Disentangling the association between genetics and functional connectivity in Mild Cognitive Impairment

Heba Elshatoury
Institute of Technology
University of Tartu
Tartu, Estonia
heba.elshatoury@ut.ee

Federica Cruciani
Dept. Computer Science
University of Verona
Verona, Italy
federica.cruciani@univr.it

Francesco Zumerle
Dept. Computer Science
University of Verona
Verona, Italy
francesco.zumerle@univr.it

Silvia F. Storti
Dept. Computer Science
University of Verona
Verona, Italy
silviafrancesca.storti@univr.it

André Altmann
Dept. Medical Physic and Biomedical Engineering
University College London
London, UK
a.altmann@ucl.ac.uk

Marco Lorenzi
Epione Research Project
Inria Sophia Antipolis
Nice, France
marco.lorenzi@inria.fr

Gholamreza Anbarjafari
Institute of Technology
University of Tartu
Tartu, Estonia
gholamreza.anbarjafari@ut.ee

Gloria Menegaz
Dept. Computer Science
University of Verona
Verona, Italy
gloria.menegaz@univr.it

Ilaria Boscolo Galazzo
Dept. Computer Science
University of Verona
Verona, Italy
ilaria.boscologalazzo@univr.it

Abstract*—Despite the increasing effort being devoted to the investigation of the link between imaging endophenotypes (IDPs) and genetic determinants (GDs) in Mild Cognitive Impairment (MCI) and Alzheimer's disease (AD), many issues remain open and deserve investigation. Among these is the role of functional connectivity (FC). This paper aims at shading some light on the topic relying on the ADNI repository. To this end, a total of 177 patients out of which 82 Mild Cognitive Impariment (MCI) and 95 healthy controls (HC) were considered. Individual FC matrices were derived after minimal pre-processing and nuisance regression, and the within/between-network connectivities were calculated to be used as IDPs. Concerning the genetics, the GDs consisted of two Polygenic Risk Scores (PRS) that have recently been proved to play a role in AD. A Partial Least Squares (PLS) model equipped with LASSO regularization was finally fitted to the data in order to assess the association between IDPs and GDs. In the first component, all the FC coefficients had the same sign, and were correlated with PRS2. Connectivities involving the dorsal attention (DAN) and frontoparietal control (CON) networks reached the highest weights, while within/between-network FC for the limbic (LIM) resulted to be less represented. Overall, the within-network FC values were less pronounced compared to the between-network ones. In the second component, most of the FC features had zero weights, with only 13 IDPs surviving. Visual (VIS) and somatomotory (SMN) showed a correlated trend, while being anti-correlated with limbic (LIM), CON and default mode network (DMN) as well as with PRS1. The between-network FC for DMN were the most represented in this second component. Our findings suggest that the two PRSs correlated with a possible pattern of aberrant within/between-network FC changes occurring in RSNs devoted to higher cognitive functions and more vulnerable in this pathology.

***Index Terms*—Resting-state functional MRI, Functional Connectivity, Partial Least Squares, Mild Cognitive Impairment, Polygenic Risk Score, Imaging genetics**

I. Introduction

Mild Cognitive Impairment (MCI) is a syndrome characterized by an accelerated cognitive decline with respect to what observed in healthy matched individuals. Despite the apparent lack of notably effects, several studies have shown that people with memory complaints and deficits have an increased risk of progression to dementia, in particular to Alzheimer's Disease (AD) [1]. While great attention has been recently given to characterising the association between genetic risk factors and brain degeneration in AD patients [2], the role of the same interaction in MCI is still less understood.

Such a process can be investigated within the Imaging Genetics (IG) framework, that is relating genetic determinants (GDs) with brain image-derived functional or structural endophenotypes (IDPs) [3].

In this context, Genome Wide Association Studies (GWAS) are generally used to identify the genetic variants associated with a given set of IDPs, relying on large cohorts of subjects. Along the same line, polygenic studies start from genetic features derived from available large scale GWAS *boiling down* the information spread across different genetic variants to one ore more scores, named Polygenic Risk Scores (PRS). These, being informative about the individual overall genetic disease

risk, allow to assess the association between genetic profiles and IDPs on smaller cohorts [4]. For AD, these scores have been shown to be associated with relevant phenotypes such as disease progression and cognitive decline [5]. However, their link with functional connectivity (FC) measures as derived from resting-state functional MRI (fMRI) data is still largely unexplored.

Blood oxygenation level dependent (BOLD) fMRI is a non-invasive method of evaluating neuronal activity in the brain either while performing a given task or at rest. Several authors have demonstrated the functional significance of the spontaneous, low frequency fluctuations (<0.1 Hz) occurring in the BOLD signal at rest and have proved the existence of spatially distinct brain areas sharing a synchronous BOLD activity, the so-called resting-state networks (RSNs) [6]. Different FC measures have been devised so far focusing either on the coherence or on the Pearson temporal correlation between time-series measured at different locations in the brain [7]. These features have been scarcely investigated in the IG framework, though could represent important biomarkers for a timely characterisation of the underlying functional modulations in neurodegenerative disorders.

In this study, we aimed at investigating the genetic influence on FC patterns in MCI patients relying on a multivariate statistical model. Our working hypothesis is that FC measures in different RSNs could reveal subtle changes induced by the onset of the disease, allowing to disentangle age-related from pathological functional degeneration and thus potentially enabling early detection of the disease fingerprints.

II. Materials and Methods

The data analysed in the current study were collected from The Alzheimer's Disease Neuroimaging Initiative (ADNI) (http://adni.loni.usc.edu/) as part of ADNI-3 phase. The selected cohort comprehended 177 subjects, including 95 healthy controls (HC, among which 52 were classified as Cognitively Normal and 43 with Significant Memory Concern) and 82 MCI (comprising 52 Early MCI, 4 MCI and 26 Late MCI, according to the ADNI database). Rs-fMRI acquisitions were performed on a 3T scanner with the following sequence parameters: TR/TE = 3000/$\sim$ 30 ms, FA = 90°, FOV = $220 \times 220 \times 163$ mm, 3.4-mm isotropic voxel size. 200 fMRI volumes were acquired in almost all subjects, with minimal variations in a small subset (e.g, 197 or 195 volumes). T1_weighted images were also available (main parameters: TR = 2300 ms, FOV = $208 \times 240 \times 256$ mm, 1-mm isotropic voxel size).

Data were preprocessed using the FMRIB Software Library (FSL version 6.0) (https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/). As minimal preprocessing, removal of the first 5 volumes, motion realignment (MCFLIRT), 4D mean intensity normalization, spatial smoothing with a 6-mm FWHM kernel and interleaved slice-timing correction were performed. A nuisance regression pipeline was then applied to regress out from the minimally preprocessed data the six motion parameters (plus their derivatives), the mean white matter/cerebro-spinal fluid (WM/CSF) signals and a linear trend component [8]. WM and CSF signals were extracted from the corresponding partial volume maps after erosion and binarisation with a threshold of 0.8. The residuals resulting from this analysis were subsequently band-pass filtered (0.01-0.08 Hz). To further eliminate motion artifacts, scrubbing was applied to remove high-motion frames as defined by exceeding 0.5 mm framewise displacement, zero-padding these volumes together with one preceding and two subsequent volumes to keep the number of data points consistent across subjects. Lastly, the preprocessed rs-fMRI images were spatially normalized to the 2-mm MNI space (non-linear registration).

The FC matrices were generated using the Schaefer functional atlas [9] with 100 parcels and 7 RSNs, namely visual (VIS), somatomotory (SMN), dorsal attention (DAN), ventral attention (VAN), limbic (LIM), frontoparietal control (CON), and default mode networks (DMN). The symmetric connectivity matrices were calculated using the Pearson correlation coefficient. In order to exploit the FC patterns in these well-known networks, summary measures representing the mean connectivity value inside a given network (within-network FC) and across edges connecting regions belonging to different networks (between-network FC) were derived from the full matrices. Within-network FC was calculated as the mean value of all the region-to-region connectivities within a specific network (e.g., DMN), while between-network FC was derived by averaging across the edges connecting a node in a network with the other nodes in the remaining networks (e.g, DMN-SMN or DMN-VIS) [10]. These operations led to 28 single FC features per subject to be used as IDPs.

Regarding the GDs, two PRSs namely PRS1 and PRS2, proposed in [11], were chosen. These scores were based on a recent GWAS study [12] and were calculated according to SNPs passing the genome-wide suggestive threshold (p=$1.0e-05$) and the p=0.5 cutoff, respectively. For further information regarding the PRS calculation please refer to [11].

Finally, a Partial Least Square (PLS) model was applied to the data to capture the association between the IDPs and GDs. The imaging features extracted represented the matrix X in the PLS model (177×28), while the genetic features composed by the two PRSs were presented as matrix Y (177×2). Before applying the model, GDs and IDPs matrices were standardised by subtracting the mean and dividing by the standard deviation. Deconfounding was also applied to remove the bias from age and gender in the X matrix. Conversely, the first five principal components of the genetic information of the whole population on which the PRS were calculated were regressed out from PRS2 only, as these represented the genetic population structures to which such PRS was highly correlated [11]. The PLS model was applied with LASSO regularization using a penalty value of 0.15. Permutation test, based on 1000 permutations of the rows of the Y matrix, was used to test the significance of the obtained eigenvalues ($p < 0.05$).

III. Results

The two components resulting from the PLS model with LASSO regularization accounted for 54% and 46% of vari-

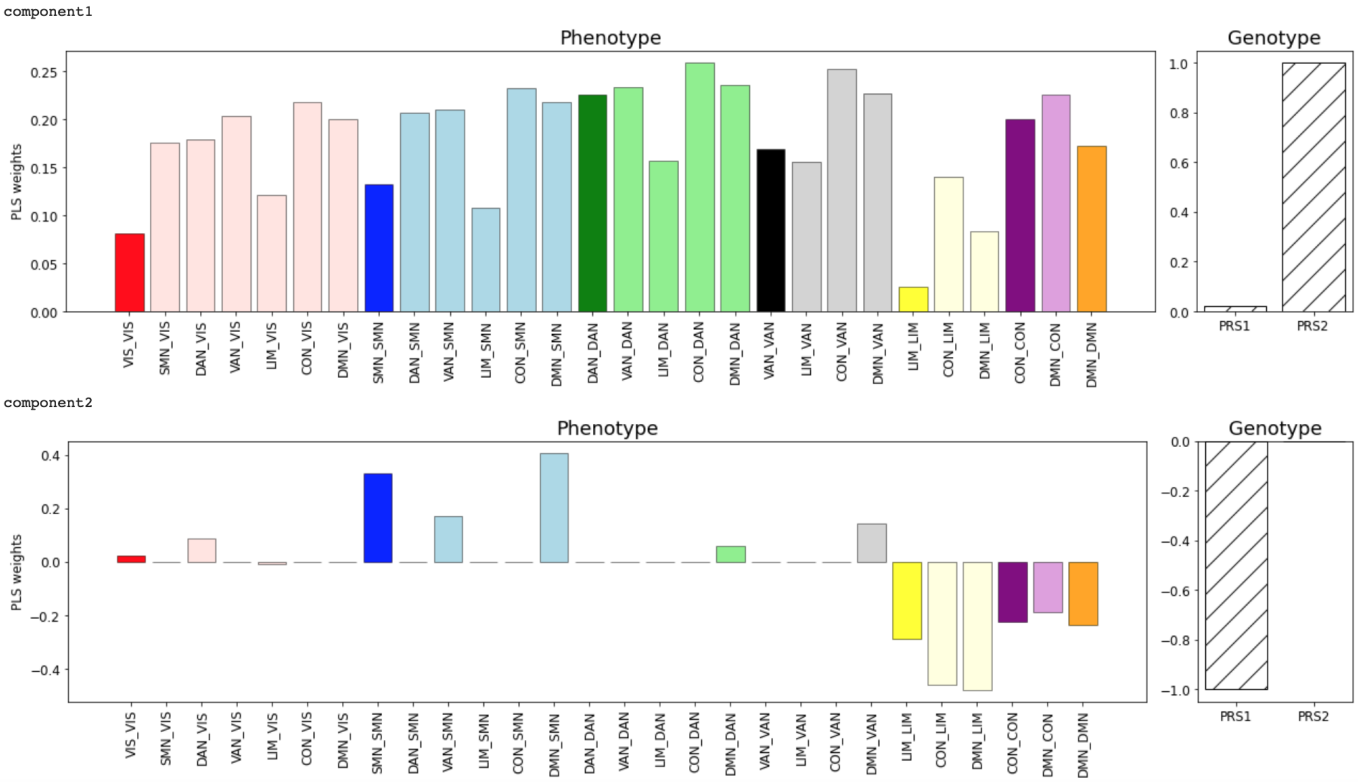

Fig. 1. PLS component's weights for imaging and genetic features. Darker colors represent the within-network connectivity while the lighter shades show the between-network connectivity.

ability of the data, respectively. The PLS weights of phenotype and genotype in both eigen-components are reported in Fig. 1. In the first component, while all the coefficients had the same sign, differences could be appreciated in terms of weights across the FC features. The within/between-network connections involving DAN generally featured the highest values, along with those encompassing the CON network. In particular, the between-network FC related to CON appeared as having the highest weights in all cases. Conversely, the connections with LIM had lower weights, especially for the intra-network one which reached the lowest value. Of note, in all cases the within-network FC values had generally lower weights compared to the between-network ones.

In the second component, most of the FC features had zero weights (or close to zero, as for LIM_VIS). The weights for the within-network FC were prominent in all cases except for DAN and VAN, and these features presented an opposite trend across networks, differently than before. Indeed, VIS and SMN showed a correlated trend, while being anti-correlated with LIM, CON and DMN. The magnitude of LIM coefficients was generally higher compared to the others, suggesting a stronger impact of this component on such FC measures. Conversely, connectivities related to the SMN, DAN and VAN networks, which reached high coefficient values in the first component, appeared to have a negligible contribution in this second one. Finally, DMN resulted to be the network with most surviving features either from within- or between-network connectivity (six out of the thirteen).

Regarding genotype variation, the PRS2 showed the highest absolute weight in the first component, while the opposite pattern was found in the second one. PRS2 was correlated with all the FC features in the first component, while in the second one PRS1 presented these correlated patterns only for a subset of connectivities involving LIM, CON and DMN. Conversely, the three features related to the SMN network with high weights (SMN_SMN, VAN_SMN, DMN_SMN) were anti-correlated with the PRS1. Finally, the permutation test proved the significance of the model with *p-value* = 0.044.

IV. Discussion

In this study we investigated the associations between neuroimaging phenotypes and genetics via joint multivariate statistical modeling in patients with MCI. The phenotypic features were presented in terms of within/between-network FC derived from rs-fMRI scans, while two PRSs were used as genetic features. These combines the effects of multiple independent risk variants into single scores, being able to capture an individual's overall genetic disease risk [11].

PLS model has been applied to maximize the covariance between the two sets of data with LASSO regularization retaining the most relevant features.

Analysis of the PLS weights showed associations between specific imaging features and one of the PRSs. In particular, all FC features were correlated with PRS2, while only LIM,

CON and DMN were correlated with PRS1 in the second component. These two PRSs have been demonstrated to be associated with APOE, clinical diagnosis, CSF-tau levels and with progressive atrophy in AD, although the second one has poorer association with traits [11]. Our results showed that the 28 FC features were differently represented in the two components and had a differential association with the two PRSs, suggesting these PRSs for AD might shape the FC fingerprints in a selective way. Interestingly, the connectivities involving DAN, VAN, CON and DMN were those featuring the highest weights in either the first or second component. These are all RSNs involved in higher cognitive functions, they comprise highly connected regions, and are characterized by an increased vulnerability compared to other networks, such as VIS or SMN, in MCI and AD patients. Previous authors argued this might depend on their particular vulnerability to amyloid deposition since the preclinical stages of dementia [13]. DMN in particular has been largely investigated in the current literature, and both within- and between network changes have been reported between HC, MCI and AD patients [14], [15]. However, several studies are going beyond DMN and have recently demonstrated aberrant inter-network changes involving those brain systems that are closely correlated and play a crucial role in higher cognitive function, underlying the central role of the interactions between RSNs in understanding MCI and AD pathology [16].

To the best of our knowledge, this was the first attempt to exploit a multivariate PLS model for linking the PRS for AD with brain FC measures in MCI subjects. In a previous work by Lorenzi and colleagues [17], PLS was used to associate brain atrophy to the complete set of SNPs from AD patients, uncovering a significant link between the TRIB3 gene and the stereotypical pattern of grey matter loss in AD. A similar approach was followed in [18], where they were able to stratify the early stages of AD in the PLS latent space by exploiting classical structural features and disease-specific biomarkers such as cerebrospinal fluid levels of t-tau, p-tau and amyloid-beta. Despite these promising preliminary results, we are aware of the small sample size which represents the main limitation of the current study and the lack of AD subjects. However, we consider the results as relevant as the within/between-network FC modulations that could be detected are in agreement with the typical neurodegeneration patterns in MCI and AD, providing evidence in favor of the suitability of PRS scores for explaining even in a selective way the genetic underpinning of such changes.

V. Conclusions

The presented PLS model with LASSO suggested a joint variation between FC and PRSs in MCI subjects. Moreover the two PRSs correlated with a possible pattern of aberrant within/between-network FC changes occurring in RSNs devoted to higher cognitive functions and more vulnerable in this pathology.

References

[1] S. Gauthier, B. Reisberg, M. Zaudig, R. C. Petersen, K. Ritchie *et al.*, "Mild cognitive impairment," *The lancet*, vol. 367, no. 9518, pp. 1262–1270, 2006.

[2] M. W. Weiner, D. P. Veitch, P. S. Aisen, L. A. Beckett, N. J. Cairns *et al.*, "2014 update of the alzheimer's disease neuroimaging initiative: a review of papers published since its inception," *Alzheimer's & dementia*, vol. 11, no. 6, pp. e1–e120, 2015.

[3] T. Steckler and G. Salvadore, "Neuroimaging as a translational tool in animal and human models of schizophrenia," in *Translational Neuroimaging*. Elsevier, 2013, pp. 195–220.

[4] I. S. Consortium, "Common polygenic variation contributes to risk of schizophrenia that overlaps with bipolar disorder," *Nature*, vol. 460, no. 7256, p. 748, 2009.

[5] M. A. Scelsi, R. R. Khan, M. Lorenzi, L. Christopher, M. D. Greicius *et al.*, "Genetic study of multimodal imaging alzheimer's disease progression score implicates novel loci," *Brain*, vol. 141, no. 7, pp. 2167–2180, 2018.

[6] B. B. Biswal, J. V. Kylen, and J. S. Hyde, "Simultaneous assessment of flow and bold signals in resting-state functional connectivity maps," *NMR in Biomedicine*, vol. 10, no. 4-5, pp. 165–170, 1997.

[7] D. M. Cole, S. M. Smith, and C. F. Beckmann, "Advances and pitfalls in the analysis and interpretation of resting-state fmri data," *Frontiers in systems neuroscience*, vol. 4, p. 8, 2010.

[8] N. Franzmeier, J. Neitzel, A. Rubinski, R. Smith, O. Strandberg *et al.*, "Functional brain architecture is associated with the rate of tau accumulation in alzheimer's disease," *Nature communications*, vol. 11, no. 1, pp. 1–17, 2020.

[9] A. Schaefer, R. Kong, E. M. Gordon, T. O. Laumann, X.-N. Zuo *et al.*, "Local-global parcellation of the human cerebral cortex from intrinsic functional connectivity mri," *Cerebral cortex*, vol. 28, no. 9, pp. 3095–3114, 2018.

[10] B. De Blasi, L. Caciagli, S. F. Storti, M. Galovic, M. Koepp *et al.*, "Noise removal in resting-state and task fmri: functional connectivity and activation maps," *Journal of Neural Engineering*, vol. 17, no. 4, p. 046040, 2020.

[11] A. Altmann, M. A. Scelsi, M. Shoai, E. de Silva, L. M. Aksman *et al.*, "A comprehensive analysis of methods for assessing polygenic burden on alzheimer's disease pathology and risk beyond apoe," *Brain communications*, vol. 2, no. 1, p. fcz047, 2020.

[12] B. W. Kunkle, B. Grenier-Boley, R. Sims, J. C. Bis, V. Damotte *et al.*, "Genetic meta-analysis of diagnosed alzheimer's disease identifies new risk loci and implicates aβ, tau, immunity and lipid processing," *Nature genetics*, vol. 51, no. 3, pp. 414–430, 2019.

[13] M. J. Grothe, S. J. Teipel, and A. D. N. Initiative, "Spatial patterns of atrophy, hypometabolism, and amyloid deposition in alzheimer's disease correspond to dissociable functional brain networks," *Human brain mapping*, vol. 37, no. 1, pp. 35–53, 2016.

[14] C. Sorg, V. Riedl, M. Mühlau, V. D. Calhoun, T. Eichele *et al.*, "Selective changes of resting-state networks in individuals at risk for alzheimer's disease," *Proceedings of the National Academy of Sciences*, vol. 104, no. 47, pp. 18 760–18 765, 2007.

[15] M. A. Binnewijzend, M. M. Schoonheim, E. Sanz-Arigita, A. M. Wink, W. M. van der Flier *et al.*, "Resting-state fmri changes in alzheimer's disease and mild cognitive impairment," *Neurobiology of aging*, vol. 33, no. 9, pp. 2018–2028, 2012.

[16] P. Wang, B. Zhou, H. Yao, Y. Zhan, Z. Zhang *et al.*, "Aberrant intra-and inter-network connectivity architectures in alzheimer's disease and mild cognitive impairment," *Scientific reports*, pp. 1–12, 2015.

[17] M. Lorenzi, A. Altmann, B. Gutman, S. Wray, C. Arber *et al.*, "Susceptibility of brain atrophy to trib3 in alzheimer's disease, evidence from functional prioritization in imaging genetics," *Proceedings of the National Academy of Sciences*, vol. 115, no. 12, pp. 3162–3167, 2018.

[18] A. Casamitjana, P. Petrone, J. L. Molinuevo, J. D. Gispert, and V. Vilaplana, "Projection to latent spaces disentangles pathological effects on brain morphology in the asymptomatic phase of alzheimer's disease," *Frontiers in neurology*, vol. 11, p. 648, 2020.