

MULTIMODAL AND LONGITUDINAL BIOIMAGING METHODS FOR CHARACTERIZING THE PROGRESSIVE COURSE OF DEMENTIA

EDITED BY: Javier Ramírez, Juan M. Górriz and Stefan Teipel
PUBLISHED IN: Frontiers in Aging Neuroscience





frontiers

Frontiers Copyright Statement

© Copyright 2007-2019 Frontiers Media SA. All rights reserved.

All content included on this site, such as text, graphics, logos, button icons, images, video/audio clips, downloads, data compilations and software, is the property of or is licensed to Frontiers Media SA ("Frontiers") or its licensees and/or subcontractors. The copyright in the text of individual articles is the property of their respective authors, subject to a license granted to Frontiers.

The compilation of articles constituting this e-book, wherever published, as well as the compilation of all other content on this site, is the exclusive property of Frontiers. For the conditions for downloading and copying of e-books from Frontiers' website, please see the Terms for Website Use. If purchasing Frontiers e-books from other websites or sources, the conditions of the website concerned apply.

Images and graphics not forming part of user-contributed materials may not be downloaded or copied without permission.

Individual articles may be downloaded and reproduced in accordance with the principles of the CC-BY licence subject to any copyright or other notices. They may not be re-sold as an e-book.

As author or other contributor you grant a CC-BY licence to others to reproduce your articles, including any graphics and third-party materials supplied by you, in accordance with the Conditions for Website Use and subject to any copyright notices which you include in connection with your articles and materials.

All copyright, and all rights therein, are protected by national and international copyright laws.

The above represents a summary only. For the full conditions see the Conditions for Authors and the Conditions for Website Use.

ISSN 1664-8714

ISBN 978-2-88945-949-0

DOI 10.3389/978-2-88945-949-0

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.

Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: researchtopics@frontiersin.org

MULTIMODAL AND LONGITUDINAL BIOIMAGING METHODS FOR CHARACTERIZING THE PROGRESSIVE COURSE OF DEMENTIA

Topic Editors:

Javier Ramírez, University of Granada, Spain

Juan M. Górriz, University of Granada, Spain

Stefan Teipel, University of Rostock and DZNE Rostock/Greifswald, Germany

Citation: Ramírez, J., Górriz, J. M., Teipel, S., eds. (2019). Multimodal and Longitudinal Bioimaging Methods for Characterizing the Progressive Course of Dementia Lausanne: Frontiers Media. doi: 10.3389/978-2-88945-949-0

Table of Contents

- 05 Editorial: Multimodal and Longitudinal Bioimaging Methods for Characterizing the Progressive Course of Dementia**
Javier Ramírez, Juan M. Górriz and Stefan Teipel
- 09 Predict Disease Progression With Reaction Rate Equation Modeling of Multimodal MRI and PET**
Li Su, Yujing Huang, Yi Wang, James Rowe and John O'Brien
- 14 Using CT Data to Improve the Quantitative Analysis of ^{18}F -FBB PET Neuroimages**
Fermín Segovia, Raquel Sánchez-Vañó, Juan M. Górriz, Javier Ramírez, Pablo Sopena-Novales, Nathalie Testart Dardel, Antonio Rodríguez-Fernández and Manuel Gómez-Río
- 24 MRI Characterizes the Progressive Course of AD and Predicts Conversion to Alzheimer's Dementia 24 Months Before Probable Diagnosis**
Christian Salvatore, Antonio Cerasa and Isabella Castiglioni for the Alzheimer's Disease Neuroimaging Initiative
- 37 Atrophy in the Thalamus but not Cerebellum is Specific for C9orf72 FTD and ALS Patients – An Atlas-Based Volumetric MRI Study**
Sonja Schönecker, Christiane Neuhofer, Markus Otto, Albert Ludolph, Jan Kassubek, Bernhard Landwehrmeyer, Sarah Anderl-Straub, Elisa Semler, Janine Diehl-Schmid, Catharina Prix, Christian Vollmar, Juan Fortea, Deutsches FTLD-Konsortium, Hans-Jürgen Huppertz, Thomas Arzberger, Dieter Edbauer, Berend Feddersen, Marianne Dieterich, Matthias L. Schroeter, Alexander E. Volk, Klaus Fließbach, Anja Schneider, Johannes Kornhuber, Manuel Maler, Johannes Prudlo, Holger Jahn, Tobias Boeckh-Behrens, Adrian Danek, Thomas Klopstock and Johannes Levin
- 48 Amyloid Associated Intermittent Network Disruptions in Cognitively Intact Older Subjects: Structural Connectivity Matters**
Susanne G. Mueller and Michael W. Weiner
- 63 Disrupted Thalamus White Matter Anatomy and Posterior Default Mode Network Effective Connectivity in Amnesic Mild Cognitive Impairment**
Thomas Alderson, Elizabeth Kehoe, Liam Maguire, Dervla Farrell, Brian Lawlor, Rose A. Kenny, Declan Lyons, Arun L. W. Bokde and Damien Coyle
- 78 Preprocessing of ^{18}F -DMFP-PET Data Based on Hidden Markov Random Fields and the Gaussian Distribution**
Fermín Segovia, Juan M. Górriz, Javier Ramírez, Francisco J. Martínez-Murcia and Diego Salas-Gonzalez
- 87 Random Forest Algorithm for the Classification of Neuroimaging Data in Alzheimer's Disease: A Systematic Review**
Alessia Sarica, Antonio Cerasa and Aldo Quattrone
- 99 Early Changes in Alpha Band Power and DMN BOLD Activity in Alzheimer's Disease: A Simultaneous Resting State EEG-fMRI Study**
Katharina Brueggen, Carmen Fiala, Christoph Berger, Sina Ochmann, Claudio Babiloni and Stefan J. Teipel

- 108** *Classifying MCI Subtypes in Community-Dwelling Elderly Using Cross-Sectional and Longitudinal MRI-Based Biomarkers*
Hao Guan, Tao Liu, Jiyang Jiang, Dacheng Tao, Jicong Zhang, Haijun Niu, Wanlin Zhu, Yilong Wang, Jian Cheng, Nicole A. Kochan, Henry Brodaty, Perminder Sachdev and Wei Wen
- 121** *Combining SPECT and Quantitative EEG Analysis for the Automated Differential Diagnosis of Disorders With Amnestic Symptoms*
Yvonne Höller, Arne C. Bathke, Andreas Uhl, Nicolas Strobl, Adelheid Lang, Jürgen Bergmann, Raffaele Nardone, Fabio Rossini, Harald Zauner, Margarita Kirschner, Amirhossein Jahanbekam, Eugen Trinkla and Wolfgang Staffen
- 136** *Visualizing Hyperactivation in Neurodegeneration Based on Prefrontal Oxygenation: A Comparative Study of Mild Alzheimer's Disease, Mild Cognitive Impairment, and Healthy Controls*
Kah Hui Yap, Wei Chun Ung, Esther G. M. Ebenezer, Nadira Nordin, Pui See Chin, Sandheep Sugathan, Sook Ching Chan, Hung Loong Yip, Masashi Kiguchi and Tong Boon Tang
- 150** *White Matter Tract Integrity in Alzheimer's Disease vs. Late Onset Bipolar Disorder and its Correlation With Systemic Inflammation and Oxidative Stress Biomarkers*
Ariadna Besga, Darya Chyzhyk, Itxaso Gonzalez-Ortega, Jon Echeveste, Marina Graña-Lecuona, Manuel Graña and Ana Gonzalez-Pinto
- 160** *Retrospective Diagnosis of Parkinsonian Syndromes Using Whole-Brain Atrophy Rates*
Carlos Guevara, Kateryna Bulatova, Wendy Soruco, Guido Gonzalez and Gonzalo A. Farías



Editorial: Multimodal and Longitudinal Bioimaging Methods for Characterizing the Progressive Course of Dementia

Javier Ramírez^{1*}, Juan M. Górriz¹ and Stefan Teipel^{2,3}

¹ Department Signal Theory, Networking and Communications, University of Granada, Granada, Spain, ² German Center for Neurodegenerative Diseases (DZNE), Rostock, Germany, ³ Department of Psychosomatic Medicine, University of Rostock and DZNE Rostock/Greifswald, Rostock, Germany

Keywords: dementia, machine learning, biomarkers, structural image analysis, functional image analysis, diagnosis and prognosis

Editorial on the Research Topic

Multimodal and Longitudinal Bioimaging Methods for Characterizing the Progressive Course of Dementia

According to the World Health Organization, in 2015 dementia affected 47 million people worldwide (or roughly 5% of the world's elderly population), a figure that is predicted to increase to 75 million in 2030 and 132 million by 2050 (World Health Organization, 2017). Dementia represents one of the major causes of disability and dependency among older people worldwide. Dementia is a broad category of mostly progressive brain diseases affecting memory, other cognitive abilities and behavior, and interfering significantly with a person's ability to maintain the activities of daily living. Alzheimer's disease (AD) is the most common cause of dementia in the elderly accounting 60–70% of cases and affects approximately 30 million individuals worldwide (Prince et al., 2013). Other major forms of dementia include vascular dementia, dementia with Lewy bodies, Parkinson's disease, frontotemporal dementia, etc.

Although new treatments are being investigated in clinical trials, no treatment to cure dementia or to alter its progressive course exists. Today, we understand that dementia appears only after a decade or more of brain degeneration (preclinical dementia) and current consensus has established the need for early recognition.

An intensive research effort is being devoted to the development of novel neuroimaging biomarkers that can provide an alert even before the cognitive decline appears. Structural and functional magnetic resonance imaging (MRI) and functional and molecular nuclear medicine neuroimaging techniques including single-photon emission computed tomography (SPECT) and positron emission tomography (PET), are widely used in combination with other blood, cerebrospinal fluid (CSF), and genetic biomarkers for early diagnosis of dementia.

Large multicenter studies are currently investigating the value of existing and novel multimodal and longitudinal neurodegeneration biomarkers. The vast amount of data available represents an opportunity for the development of more accurate statistical models of neurodegeneration enabling the early recognition as well as the characterization of the progressive course of dementia.

The aim of the Research Topic “Multimodal and Longitudinal Bioimaging Methods for Characterizing the Progressive Course of Dementia,” published in *Frontiers in Aging Neuroscience*, was to present the current state of the art in the theory and practice of multimodal and longitudinal neuroimaging analysis approaches for characterizing the progressive course of dementia. The Research Topic features 14 research articles. Most of the contributions analyzed disease progression and the relationships among underlying pathological changes.

OPEN ACCESS

Edited and reviewed by:

Thomas Wisniewski,
New York University School of
Medicine, United States

*Correspondence:

Javier Ramírez
javierrp@ugr.es

Received: 12 January 2019

Accepted: 18 February 2019

Published: 14 March 2019

Citation:

Ramírez J, Górriz JM and Teipel S
(2019) Editorial: Multimodal and
Longitudinal Bioimaging Methods for
Characterizing the Progressive Course
of Dementia.
Front. Aging Neurosci. 11:45.
doi: 10.3389/fnagi.2019.00045

Differentiating between Parkinson's disease (PD) and atypical parkinsonian syndromes (APS) is still a challenge, specially at early stages when the patients show similar symptoms. During last years, several computational approaches have been proposed in order to improve the diagnosis of PD, but their accuracy is still limited (Segovia et al., 2015, 2017). The first paper of the Research Topic is devoted to the development of analysis methods for diagnosis of idiopathic Parkinson's disease (IPD), multiple system atrophy (MSA), and progressive supranuclear palsy (PSP) (Guevara et al.). Ten healthy controls, 20 IPD, 39 PSP, and 41 MSA patients were studied using MRI and Structural Imaging Evaluation with Normalization of Atrophy (SIENA) (Smith et al., 2002).

Bipolar disorders such as the Late Onset Bipolar Disorder (LOBD) is often difficult to be differentiated from neurodegenerative dementias due to common cognitive and behavioral impairment symptoms. In a multimodal study Besga et al. determined differences in white matter (WM) tract integrity between AD and LOBD cases, and their correlation with systemic inflammatory, neurotrophic factors, and oxidative stress blood plasma biomarkers. Differences in WM tract integrity reflected greater behavioral and mood clinical features of LOBD and together with alterations of neuroinflammatory blood markers, different impact of neuroinflammation in both diseases.

The paper by Yap et al. focuses on visualization of hyperactivation in neurodegeneration based on prefrontal oxygenation showing a comparative study of mild AD dementia, mild cognitive impairment, and healthy controls. Functional near-infrared spectroscopy (fNIRS) signals were analyzed together with a semantic verbal fluency task (SVFT) to investigate any compensation exhibited by the prefrontal cortex (PFC). It was shown that the task-elicited hyperactivation in MCI might reflect the presence of compensatory mechanisms, and hypoactivation in mild AD dementia could reflect an inability to compensate.

Several works have suggested that multimodal data analysis has the potential to improve the diagnosis of dementia (Ortiz et al., 2018). The paper by Höller et al. showed that combining quantitative markers from SPECT and EEG increased discrimination of MCI and AD cases from people with depression, and that the resulting diagnostic accuracies were higher than the diagnostic accuracy of each single modality alone.

The paper by Guan et al. addresses the development of MCI subtype classification techniques to enable early intervention with targeted treatment. A sample of 184 community-dwelling individuals (aged 73–85 years) was analyzed and cortical surface based measurements were computed from longitudinal and cross-sectional MRI scans. Their results using feature selection and a voting classifier suggested that longitudinal features were not superior to the cross-sectional features for MCI subtype classifications.

The paper by Sarica et al. provides a systematic review of random forests (Ramírez et al., 2009, 2010, 2018) as an enabling machine learning technique for automatic early diagnosis and prognosis of AD using single and multi-modal neuroimaging data.

Emerging imaging modalities are also covered in this Research Topic. Simultaneous EEG-fMRI acquisitions allow combining the spatial resolution of fMRI with the temporal resolution of EEG. The paper by Brueggen et al. carried out a study of simultaneous fMRI-EEG acquisitions in a sample of AD patients and controls and showed a reduced positive association between alpha band power and BOLD fluctuations in AD patients, compared to the control subjects. ^{18}F -DMFP-PET is a neuroimaging modality used to diagnose Parkinson's disease (PD) by examining postsynaptic dopamine D2/3 receptors. Segovia et al. proposed a novel methodology to preprocess ^{18}F -DMFP-PET data that improves the accuracy of computer aided diagnosis systems for PD. PET data were segmented into 4 maps according to the intensity and the neighborhood of the voxels using an algorithm based on Hidden Markov Random Field. Then, the maps were individually normalized so that the shape of their histograms could be modeled by a template Gaussian distribution. The results outperformed those reported by previous approaches.

The article by Alderson et al. used a multimodal approach to assess white matter integrity between thalamus and default mode network (DMN) components and associated effective connectivity in healthy controls (HCs) relative to aMCI patients. Their methodology enabled the DMN of each subject to be identified using independent component analysis (ICA) and resting state effective connectivity that was calculated between thalamus and DMN nodes. Significant changes in the diffusivity metrics of thalamic white matter projection tracts to hippocampus, posterior cingulate cortex and lateral inferior parietal lobe were identified.

Based on the notion that amyloid may induce neuronal network hypersynchrony in early AD stages the work by Mueller and Weiner developed a graph and cluster analysis on a sample of Florbetapir-F18 PET and task-free 3T functional and structural MRI and found distinct pattern of hypersynchrony with underlying white matter connectivity in amyloid positive vs. negative cognitively normal older subjects.

Mutation carriers may exhibit distinct neuropathological features of neurodegenerative diseases. As an example, patients with frontotemporal dementia (FTD) or amyotrophic lateral sclerosis (ALS) due to a C9orf72 mutation are characterized by two distinct types of characteristic protein depositions. The study by Schönecker et al. aimed to determine if mutation carriers showed an enhanced degree of thalamic and cerebellar atrophy compared to sporadic FTD and ALS patients or healthy controls.

The paper by Salvatore et al. analyzed progression of AD using a machine learning method in a cohort of 200 subjects obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI). Subjects were followed-up for 24 months, and grouped as AD, progressive-MCI to AD, stable-MCI, and cognitively normal (CN). Structural T1-weighted MRI and neuropsychological measures were used to train a classifier to distinguish mild-AD patients (AD + progressive MCI) from subjects with a benign cognitive course (stable-MCI + CN). Principal component analysis (PCA) and partial Least squares (PLS) were used as feature extraction methods similar to previous studies (López et al., 2011; Segovia et al., 2012; Khedher et al., 2015).

^{18}F -labeled amyloid tracers have been approved with similar efficacy to PIB and longer half-life: ^{18}F -florbetapir in 2012, ^{18}F -flutemetamol in 2013 and ^{18}F -florbetaben (FBB) in 2014. Based on the broad availability of PET-CT scanners, the work by Segovia et al. proposed to include in the analysis the information about gray matter neurodegeneration provided by CT images in order to improve the diagnosis of AD. Specifically, standardized uptake values (SUVs) from ^{18}F -FBB PET data were obtained using only voxels belonging to gray matter in CT images. The results suggested that SUVs calculated according to the proposed method allowed AD and non-AD subjects to be more accurately differentiated. This agrees with previous studies on the use of structural MRI scans to correct amyloid PET data for spill out effect of signal from gray matter to CSF and for spill in effect from white to gray matter (Gonzalez-Escamilla et al., 2017).

Previous works have shown that beta-amyloid, tau, neuroinflammation and neurodegeneration all play a significant role in the etiology of Alzheimer's disease (AD) (Lehmann et al., 2013). In Su et al. a novel computational modeling approach for multimodal MRI and PET inspired by reaction rate equation in chemical kinetics is proposed to investigate the progression of AD and relationships among underlying pathological changes. The study is motivated by the fact that the relationship between them is often unclear, mainly because the time scale associated to dementia generally exceeds the one of other studies and the challenge of observing the ordering of the pathological changes during the progression of the disease.

In summary, the Research Topic provides a transnosological, transdisciplinary view on current developments in neuroimaging techniques and their application to neurodegenerative

dementias. It depicts a dynamic landscape of emerging acquisition and analyses techniques that share, however, three key features:

- The combination of modalities within and beyond imaging techniques to improve group separation and modeling of longitudinal decline
- The combination of data driven and model driven analysis approaches to get the best from both worlds: full exploitation of available data and top down restriction of the outcome space based on a priori assumptions on disease pathogenesis
- Related to these features, the ability to test assumptions of underlying disease mechanisms in combinatorial multimodal imaging analyses, crossing even the boundaries of single disease entities.

Such highly interdisciplinary approach serves as blueprint not only for future research in neurodegenerative dementias but in other neuropsychiatric diseases as well.

AUTHOR CONTRIBUTIONS

JR, JG, and ST equally participate in the preparation of this manuscript.

ACKNOWLEDGMENTS

This work was partly supported by the MINECO/FEDER under the TEC2015-64718-R project and the Consejera de Economía, Innovación, Ciencia y Empleo (Junta de Andalucía, Spain) under the Excellence Project P11-TIC-7103.

REFERENCES

- Gonzalez-Escamilla, G., Lange, C., Teipel, S., Buchert, R., Grothe, M. J., and Alzheimer's Disease Neuroimaging Initiative (2017). PETPVE12: an SPM toolbox for Partial Volume Effects correction in brain PET - Application to amyloid imaging with AV45-PET. *Neuroimage* 147, 669–677. doi: 10.1016/j.neuroimage.2016.12.077
- Khedher, L., Ramírez, J., Górriz, J., Brahim, A., and Segovia, F. (2015). Early diagnosis of Alzheimers disease based on partial least squares, principal component analysis and support vector machine using segmented MRI images. *Neurocomputing* 151, 139–150. doi: 10.1016/j.neucom.2014.09.072
- Lehmann, M., Ghosh, P. M., Madison, C., Laforce, R., Corbetta-Rastelli, C., Weiner, M. W., et al. (2013). Diverging patterns of amyloid deposition and hypometabolism in clinical variants of probable Alzheimer's disease. *Brain* 136, 844–858. doi: 10.1093/brain/awt327
- López, M., Ramírez, J., Górriz, J., Álvarez, I., Salas-Gonzalez, D., Segovia, F., et al. (2011). Principal component analysis-based techniques and supervised classification schemes for the early detection of Alzheimer's disease. *Neurocomputing* 74, 1260–1271. doi: 10.1016/j.neucom.2010.06.025
- Ortiz, A., Lozano, F., Górriz, J. M., Ramírez, J., Martínez Murcia, F. J., and Alzheimer's Disease Neuroimaging Initiative (2018). Discriminative sparse features for Alzheimer's disease diagnosis using multimodal image data. *Curr. Alzheimer Res.* 15, 67–79. doi: 10.2174/1567205014666170922101135
- Prince, M. J., Prina, M., and Guerchet, M. (2013). *World Alzheimer Report 2013. Journey of Caring. An Analysis of Long-term Care for Dementia*. Alzheimer's Disease International.
- Ramírez, J., Górriz, J. M., Chaves, R., López, M., Salas-Gonzalez, D., Alvarez, I., et al. (2009). SPECT image classification using random forests. *Electron. Lett.* 45, 604–605. doi: 10.1049/el.2009.1111
- Ramírez, J., Górriz, J. M., Ortiz, A., Martínez-Murcia, F. J., Segovia, F., Salas-Gonzalez, D., et al. (2018). Ensemble of random forests One vs. Rest classifiers for MCI and AD prediction using ANOVA cortical and subcortical feature selection and partial least squares. *J. Neurosci. Methods* 302, 47–57. doi: 10.1016/j.jneumeth.2017.12.005
- Ramírez, J., Górriz, J. M., Segovia, F., Chaves, R., Salas-Gonzalez, D., López, M., et al. (2010). Computer aided diagnosis system for the Alzheimer's disease based on partial least squares and random forest SPECT image classification. *Neurosci. Lett.* 472, 99–103. doi: 10.1016/j.neulet.2010.01.056
- Segovia, F., Górriz, J. M., Ramírez, J., Alvarez, I., Jimenez-Hoyuela, J. M., and Ortega, S. J. (2012). Improved parkinsonism diagnosis using a partial least squares based approach. *Med. Phys.* 39, 4395–4403. doi: 10.1118/1.4730289
- Segovia, F., Górriz, J. M., Ramírez, J., Martínez-Murcia, F. J., Levin, J., Schuberth, M., et al. (2017). Multivariate analysis of ^{18}F -dmfp pet data to assist the diagnosis of parkinsonism. *Front. Neuroinformatics* 11:23. doi: 10.3389/fninf.2017.00023

- Segovia, F., Illán, I. A., Górriz, J. M., Ramírez, J., Rominger, A., and Levin, J. (2015). Distinguishing parkinson's disease from atypical parkinsonian syndromes using pet data and a computer system based on support vector machines and bayesian networks. *Front. Comput. Neurosci.* 9:137. doi: 10.3389/fncom.2015.00137
- Smith, S. M., Zhang, Y., Jenkinson, M., Chen, J., Matthews, P. M., Federico, A., et al. (2002). Accurate, robust, and automated longitudinal and cross-sectional brain change analysis. *Neuroimage* 17, 479–489. doi: 10.1006/nimg.2002.1040
- World Health Organization (2017). *Global Action Plan on the Public Health Response to Dementia 2017 - 2025*. WHO.

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.

Copyright © 2019 Ramírez, Górriz and Teipel. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Predict Disease Progression With Reaction Rate Equation Modeling of Multimodal MRI and PET

Li Su^{1,2*}, Yujing Huang¹, Yi Wang¹, James Rowe^{3†} and John O'Brien^{1†}

¹ Department of Psychiatry, University of Cambridge, Cambridge, United Kingdom, ² China-UK Centre for Cognition and Ageing Research, Faculty of Psychology, Southwest University, Chongqing, China, ³ Department of Clinical Neurosciences, University of Cambridge, Cambridge, United Kingdom

Neurodegenerative dementia often has multiple types of underlying pathology, for example, beta-amyloid, misfolded tau, chronic neuroinflammation and neurodegeneration may coexist in Alzheimer's disease. However, the relationship between them is often unclear, in other words, whether one pathology is upstream or downstream of others can be very difficult to investigate directly. This is partly because the underlying pathology in dementia may precede detectable symptoms by several years if not decades. The time scale associated with disease progression in dementia generally exceeds that in conventional longitudinal imaging studies in humans, so it is difficult to directly observe the temporal ordering of different pathologies. Also, animal studies are not always transferable to patients due to obvious differences between the two systems. To investigate the disease progression and relationships among underlying pathological changes, we propose a novel computational modeling approach for multimodal MRI and PET inspired by *reaction rate equation* in chemical kinetics. We also discuss the possibility and prerequisites to use cross-sectional data to generate preliminary hypothesis for future longitudinal studies. It has been shown that the rate of change in some biomarkers can be approximated by the average trajectory across patients at different stages of disease severity in cross-sectional studies. The relationship modeled in our approach is akin to that in the control theory, and can be assessed by demonstrating that the presence of one disease related biomarker predicts dynamics in another. We argue that the proposed framework has important implications for trials targeting different pathologies in dementia.

Keywords: MRI, PET, computational modeling, disease progression, AD, dementia

OPEN ACCESS

Edited by:

Juan Manuel Gorriz,
Universidad de Granada, Spain

Reviewed by:

Johannes Levin,
Ludwig-Maximilians-Universität
München, Germany

Neil P. Oxtoby,
University College London,
United Kingdom

*Correspondence:

Li Su
ls514@cam.ac.uk

† Joint last authorship

Received: 07 March 2018

Accepted: 14 September 2018

Published: 08 October 2018

Citation:

Su L, Huang Y, Wang Y, Rowe J
and O'Brien J (2018) Predict Disease
Progression With Reaction Rate
Equation Modeling of Multimodal
MRI and PET.
Front. Aging Neurosci. 10:306.
doi: 10.3389/fnagi.2018.00306

INTRODUCTION

Previous studies have shown that beta-amyloid, tau, neuroinflammation and neurodegeneration all play a significant role in the etiology of Alzheimer's disease (AD), but little is known about their relationships (Edison et al., 2008; Lehmann et al., 2013). In particular, whether one type of pathology is the upstream or downstream event to others has significant impact on future trials appropriately targeting them at the right point in the disease course (Jack et al., 2013). In addition, systematically determining potential treatment

targets in diseases with multiple interacting pathologies has strategic importance for effective treatments. In the case of AD, beta amyloid has been regarded as an early event in the disease progression therefore making it one of the potential targets (Jack et al., 2013). However, the failures of several recent trials of anti-amyloid therapy (Le Couteur et al., 2016) may arguably be caused by the drugs were given too late in the disease course to be effective. The co-existence of other interacting pathologies might also reduce the efficacy of those anti-amyloid drugs. So, investigating the relationship between multiple pathologies and their associated imaging biomarkers in dementia and determining alternative or complementary treatment targets are necessary.

The relationship among more than one type of pathology has primarily studied in animal models of AD, however, due to the differences in techniques, the findings from those animal research were largely inconsistent (Yoshiyama et al., 2007). In one study, activated microglia was found to facilitate the propagation of misfolded tau in mouse brains (Wes et al., 2014), and after depletion of microglia in the mouse brain, the spread of tau from the entorhinal cortex to the dentate gyrus was significantly decrease (Asai et al., 2015). This evidence points to a causal effect of neuroinflammation on the phosphorylation or propagation of tau as microglia activations seems both sufficient and necessary for tau phosphorylation and its *trans*-synaptic propagation. Although some studies showing microglia activation preceded tangle formation in P301S transgenic mice with overexpressed mutant human tau, opposite pattern was found in Cx3cr1 mice with tau deficiency that shows tau phosphorylation without significant microglia activation and reduced neuroinflammation (Yoshiyama et al., 2007). In addition, the obvious differences between humans and animals limit the ability to translate findings from the modal systems to human patients. To see how different types of pathology interact in humans, it has required a multimodal imaging study with PET and MRI in the same cohort of participants to reveal the potential influences among them. Recently such data is emerging, but the analysis and the inference frameworks still lag behind in characterizing multimodal and longitudinal imaging data in patients with dementia.

In neurodegenerative dementia such as AD, the development of underlying pathology takes several years if not decades before any detectable symptoms occur (Jack et al., 2013), so it is challenging to study in humans using conventional longitudinal design. This is because it is costly and difficult to follow large cohort of healthy participants free from AD pathologies over many decades with only very small proportion of them eventually develops dementia. For imaging studies, the MR scanner will also unavoidably change over time, making the data less comparable at different time points. As a result, existing longitudinal human imaging data only tracks a relatively short period (e.g., several years) within the evolution of the disease in patients with dementia (Ishiki et al., 2015). In addition, different biomarkers may have different sensitivity to the underlying pathology. So comparing biomarkers obtained from multimodal imaging is nontrivial. With the absence of suitable longitudinal data tracking the long-term evolution of dementia in humans and the inconsistent evidence from animal models, studying interaction

among factors of dementia may sometimes rely on cross-sectional data and by modeling the relationship within clinical populations representing different severities and stages of disease progression. Thus, ideal analysis methods for longitudinal data must also consider cross-sectional data to be widely applicable in dementia research.

Here, we proposed a novel computational modeling approach based on *reaction rate equation* modeling in chemical kinetics to infer relationship between more than one types of imaging biomarker in dementia. The specific type of relationship in this model was defined in a control theory sense (Friston et al., 2016) meaning whether the presence of one disease related pathological process (e.g., tau) in the past predicts to the dynamics in another (e.g., beta-amyloid, microglia activation, and neurodegeneration) expressed using differential equations (Yang et al., 2011; Villemagne et al., 2013; Young et al., 2014; Budgeon et al., 2017; Lturria-Medina et al., 2017; Oxtoby et al., 2018).

THE LONGITUDINAL MODEL

With the advances of imaging technology, many types of pathology can be measured *in vivo* using multimodal MRI and multi-tracer PET within the same cohort (Passamonti et al., 2016, 2018; Su et al., 2016). For example PET data is often in the form of binding potential or SUVR that are proxies for the concentration of a substance, e.g., some neurotoxic proteins. MRI data is often in the form of gray matter volume as well as cortical thickness. For simplicity, we will illustrate the approach by the interaction between two substances measured from PET imaging, each associated with a specific pathology. However, this model can easily be extended to MRI and between PET and MRI.

Specifically, if one substance BP_1 (related to one pathology) is the upstream species of another substance BP_2 (related to a different pathology) in a biochemical reaction, we can express this as Equation 1 (which is called balance equation).

$$aBP_1 + b_i \Sigma B_i \rightarrow mBP_2 + n_j \Sigma C_j \quad (1)$$

Where B_i and C_j represent a set of unknown substances involved in this process and a , b_i , m and n_j represent a set of unknown coefficients. Also, i and j are indices of these unknown substances of which the number of them are also unknown. Here, we do not assume this process is a single step reaction nor substance represented by BP_1 directly turns into BP_2 . It can be seen that Equation 1 describes a specific form of relationship, i.e., BP_1 is an up-stream event of BP_2 instead of the vice versa. This model is difficult to estimate because it contains too many free parameters that we cannot evaluate empirically in patients. It will be discussed later that Equation 1 remains useful to conceptualize the relationship between multiple types of pathology in the disease.

To reduce the number of free parameters in the model, we will use chemical kinetics to calculate the speed of the reaction. Here, we do not need to estimate the concentration of all products in C_j if we can instead measure the speed at which BP_2 is produced because it will be perfectly correlated with the speed at which

the reaction happens. In chemical kinetics, the speed at which a reaction happens is expressed using the reaction rate equation (Equation 2).

$$\frac{d[BP_2]}{dt} = -\frac{d[BP_1]}{dt} = k[BP_1]^x \Pi[B_i]^y \quad (2)$$

Where k is a reaction rate constant, x and y are unknown reaction orders, and the concentrations $[BP_1]$ and $[BP_2]$ can be measured by PET from human participants, example of which are $[^{18}\text{F}]\text{AV1451}$ for tau, $[^{11}\text{C}]\text{PiB}$ for beta-amyloid and $[^{11}\text{C}]\text{PK11195}$ for activated microglia. $[A]$ denotes the concentration of substance A while $[A]_0$ represents the initial value of the concentration at a predefined baseline time point. In most diseases, the initial concentration of disease related substance $[BP_1]_0$ is significantly lower compared with other substances $[B_i]_0$ that were already in the brains of healthy subjects (i.e., $[B_i]_0 \gg [BP_1]_0$), we can apply the pseudo 1st order approximation in chemical kinetics and simplify Equation 2 to Equation 3 that only contains the concentration $[BP_1]$. It should be noted that the validity of this assumption is also depending on the selection of the initial state. For AD, the MCI stage is a relatively reasonable and clinical defined initial state for the disease process, however, as new method for the early detection of AD emerges, a more accurate initial state could be defined in the future. Thus, for the longitudinal model, an ideal baseline state is at the very early point of the diseases. As previously mentioned, longitudinal data with such suitable baseline is still lacking.

$$\frac{d[BP_2]}{dt} = k[BP_1]^x \Pi[B_i]_0^y = k'[BP_1]^x \quad (3)$$

Where k' is a constant because we assume that the initial concentrations of $[B_i]_0$ are also constants. Finally, we apply natural logarithm on both sides of the non-linear Equation 3, which is more difficult to model and evaluate, resulting an asymmetric linear model (Equation 4), which can be statistically tested using linear regression as we explained in subsequent sections.

$$\ln\left(\frac{d[BP_2]}{dt}\right) = \ln(k'[BP_1]^x) = \ln(k') + x \ln([BP_1]) \quad (4)$$

THE CROSS-SECTIONAL MODEL

As previously discussed, the progression of dementia often takes years or decades, thus conventional longitudinal data are either unavailable or unable to capture the full dynamic changes and temporal ordering in underlying pathologies. Thus, short-term follow-up data was often used to fit the common biomarker trajectory building up a complete picture of the population time course over a longer term (Budgeon et al., 2017). Although the longitudinal model is still more powerful and likely to be more accurate in revealing relationship between biomarkers, the method proposed here should be considered with respect to cross-sectional data when longitudinal data is not available see similar approached used by Young et al. (2014). This allows us to generate hypothesis from existing cross-sectional studies, and then to test it with longitudinal data in the future.

In some cases, the average rate of change in tau pathology over time has been shown to be consistent with their rate of change across individual subjects with different disease severity or disease progression score (Ishiki et al., 2015). In the cross-sectional model under the above condition, we replace the total derivative with respect to time in Equation 4 by partial derivative with respect to an appropriate measurement of disease severity or cognitive functioning (denoted by τ); see Equation 5.

$$\ln\left(\frac{\partial[BP_2]}{\partial\tau}\right) = \ln(k') + x \ln([BP_1]) \quad (5)$$

In this model, we approximate the rates of longitudinal change in regional PET binding and GM density for a cohort of patients by the slope parameters with respect for the disease severity score or cognition derived from a multiple linear regression model. Here, we assumed that the pathological changed linearly with the disease severity of cognitive measure. This assumption although not true in the absolute sense, it avoids over-fitting the noise when the sample size is limited, for example in most PET studies. In order to control other known confounding factors, one can include subject's age, sex, and years of education as covariates in the regression analysis.

STATISTICAL ANALYSIS

A critical step for computational modeling is empirical validation, in other words, whether the proposed model explains human data. So, one should fit the reaction rate model to the imaging data, either longitudinal or cross-sectional. As the model is a linear equation, the "goodness of fit" can be statistically evaluated by linear correlation. In other words, if we hypothesize that one type of pathology is the upstream event of another under the assumption and formulation of this model, baseline level of the former should be correlated with the rate of changes in the latter either over time for longitudinal data or across different degrees of dementia severities for cross-sectional data. This approach can be extended in several ways. In addition, we can also apply the modeling to the clinical and cognitive data. Although they cannot be directly formulated within the context of chemical reactions, the quantitative method does allow inference to be drawn between these metrics and brain imaging or other biomarkers.

The statistical tests for longitudinal and cross-sectional data may be different. In the longitudinal case, the dynamic model can be fitted to each individual subject's time course data, hence the group statistics can be computed using a random effect model across multiple subjects after the model fitting at the individual level for each brain area. However, in the case of cross-sectional data or longitudinal data with short-term follow-up, the rate of biomarker changes can only be estimated for the cohort as a whole. Thus, the test for the fitness of the model cannot be done using the methods for longitudinal data. Instead, it can be done across different brain regions using repeated measure statistics, and the inference can only be drawn for the whole brain. In addition, for the cross-sectional model, including cognitive or clinical data is more difficult than the longitudinal design because

the rates of the biomarker changes are estimated for the group rather than for the individuals.

Another challenging issue in computational modeling is model comparison. Different methods have been proposed to account for this issue, such as Bayesian model selection, which gained increasing popularity in imaging analysis and generative modeling in computational neuroscience (Wasserman, 2000). These sophisticated methods deal with difference in model complexity, i.e., the simpler model is the superior model compared with a more complex model if both models can explain the data equally well. This can be understood by the intuitive Occam's Razor principle, i.e., preferring the most parsimonious model explaining the same variance. However, in our formalism, all models have not only the same number of parameters but also the same analytical form (Equation 4 or 5). As the complexities between alternative models are identical, model comparison is trivial and can be done by simply comparing the goodness of fit.

DISCUSSION

It is widely accepted that longitudinal imaging is very important for the understanding of disease progression, staging pathology, differential diagnosis, and determining prognosis during clinical trials. Multimodal imaging including structural MRI, DTI, functional MRI and multiple tracers PET has also gained the attention because it allows complex etiologies behind dementing diseases to be studied (Mak et al., 2014). However, the mainstream analysis methods are still limited in their capacity to handle longitudinal data and to systematically relate or combine data from multiple imaging modalities. Moreover, mechanistic interference on the interactions among different pathologies and their associated biomarkers cannot be reliably drawn for the following reasons. First, the majority of imaging data on dementia is either cross-sectional or only tracks the disease progression for a few years and often during the relatively late stages, e.g., after significant cognitive impairments and brain damages have occurred. Such short-term follow-up data may be difficult to reliably capture critical events during the disease course, in particular during the pre-symptomatic stages. Second, existing imaging data is often acquired from single site with relatively small heterogeneous samples. The individual differences among patients are often dramatic because the variation in clinical diagnosis criteria and age as well as the possibility of mixed pathologies or even misdiagnosis between dementia sub-types.

To analyze and model longitudinal and multimodal imaging data with these limitations, several methods have been proposed. For example, imaging biomarkers of pathology measured at different time points from different individuals can be mapped onto a hypothetical time axis representing 'time-to-dementia' (Bateman et al., 2012). This approach allows us to normalize different biomarkers from the heterogeneous sample to a single continuous dimension so that temporal ordering can be inferred. Other approaches use data-driven methods to model biomarkers from different data points as discrete events in the disease course. The temporal ordering of events can then be inferred based on the co-occurrence

between each pair of different events and Markov chain Monte Carlo methods (Young et al., 2014; Oxtoby et al., 2018).

One advantage of our approach is that the hypothesis about the relationship between different imaging biomarkers are expressed explicitly as chemical reactions, i.e., using the balance equation (Equation 1). Hence, underlying assumptions have to be made explicit, an advantage of most computational models. In this formalism, it is apparent that the temporal ordering between different biomarkers cannot be inferred from the quantity difference between each modeled substance alone. For example, a greater concentration or binding potential in BP_1 than in BP_2 does not always imply that pathology related to BP_1 is the upstream event of BP_2 . This is because that the coefficients such as a and m in Equation 1 are unknown, so it is possible that a small quantity of one substance at upstream results in a larger quantity of another downstream, i.e., when $a > m$ in Equation 1. This may give the false inference that the downstream event related to BP_2 is preceding the upstream event related to BP_1 . By the same token, the temporal ordering of the biomarkers cannot be solely determined from the spatial extent which is another common way to measure quantity in neuroimaging.

Finally, we argue that this dynamic perspective in modeling biomarkers may be extended in order to be applied to not only neurodegenerative diseases but also neurodevelopmental conditions such as ADHD and autism spectrum disorders as well as normal development and aging. However, the limitation of our methods is that it still assumes stationarity, i.e., the rate of change in biomarkers (i.e., the 2nd order derivatives) does not change over time. Future developments are needed to capture the nonlinearity of the disease progression and heterogeneity of the samples. Last but not least, this highly novel approach requires extensive empirical validations using suitable longitudinal and cross-sectional data from different diseases and ideally from cohorts of different age ranges.

AUTHOR CONTRIBUTIONS

LS designed the study, developed the computational model, and wrote the manuscript. YH and YW co-developed the computational model. JR and JO'B co-designed the study, obtained funding for the project, and assisted with writing of the manuscript.

FUNDING

The study was funded by the National Institute for Health Research (NIHR) Biomedical Research Centre and Biomedical Research Unit in Dementia based at Cambridge University Hospitals NHS Foundation Trust and the University of Cambridge. We thank the support from Alzheimer's Research United Kingdom (ARUK-SRF2017B-1), Addenbrooke's Charitable Trust and the Lewy Body Society. JR is supported by the Wellcome Trust (103838). The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health.

REFERENCES

- Asai, H., Ikezu, S., Tsunoda, S., Medalla, M., Luebke, J., Haydar, T., et al. (2015). Depletion of microglia and inhibition of exosome synthesis halt tau propagation. *Nat. Neurosci.* 18, 1584–1593. doi: 10.1038/nn.4132
- Bateman, R. J., Xiong, C., Benzinger, T. L., Fagan, A. M., Goate, A., Fox, N. C., et al. (2012). Clinical and biomarker changes in dominantly inherited Alzheimer's disease. *N. Engl. J. Med.* 367, 795–804. doi: 10.1056/NEJMoa1202753
- Budgeon, C. A., Murray, K., Turlach, B. A., Baker, S., Villemagne, V. L., Burnham, S. C., et al. (2017). Constructing longitudinal disease progression curves using sparse, short-term individual data with an application to Alzheimer's disease. *Stat. Med.* 36, 2720–2734. doi: 10.1002/sim.7300
- Edison, P., Archer, H. A., Gerhard, A., Hinz, R., Pavese, N., Turkheimer, F. E., et al. (2008). Microglia, amyloid and cognition in Alzheimer's disease: an 11CPRK11195-PET and 11CPIB-PET study. *Neurobiol. Dis.* 32, 412–419. doi: 10.1016/j.nbd.2008.08.001
- Friston, K., Litvak, V., Oswal, A., Razi, A., Stephan, K. E., van Wijk, B. C., et al. (2016). Bayesian model reduction and empirical bayes for group (DCM) studies. *Neuroimage* 128, 413–431. doi: 10.1016/j.neuroimage.2015.11.015
- Ishiki, A., Okamura, N., Furukawa, K., Furumoto, S., Harada, R., Tomita, N., et al. (2015). Longitudinal assessment of tau pathology in patients with Alzheimer's disease using [18F]THK-5117 positron emission tomography. *PLoS One* 10:e0140311. doi: 10.1371/journal.pone.0140311
- Jack, C. R., Knopman, D. S., Jagust, W. J., Petersen, R. C., Weiner, M. W., Aisen, P. S., et al. (2013). Update on hypothetical model of Alzheimer's disease biomarkers. *Lancet Neurol.* 12, 207–216. doi: 10.1016/S1474-4422(12)70291-0
- Le Couteur, D. G., Hunter, S., and Brayne, C. (2016). Solanezumab and the amyloid hypothesis for Alzheimer's disease. *BMJ* 355:i6771. doi: 10.1136/bmj.i6771
- Lehmann, M., Ghosh, P. M., Madison, C., Laforce, R. Jr., Corbetta-Rastelli, C., Weiner, M. W., et al. (2013). Diverging patterns of amyloid deposition and hypometabolism in clinical variants of probable Alzheimer's disease. *Brain* 136, 844–858. doi: 10.1093/brain/aww327
- Lturría-Medina, Y., Carbonell, F. M., Sotero, R. C., Chouinard-Decorte, F., Evans, A. C., Alzheimer's Disease Neuroimaging Initiative (2017). Multifactorial causal model of brain (dis)organization and therapeutic intervention: application to Alzheimer's disease. *Neuroimage* 152, 60–77. doi: 10.1016/j.neuroimage.2017.02.058
- Mak, E., Su, L., Williams, G. B., and O'Brien, J. T. (2014). Neuroimaging characteristics of dementia with lewy bodies. *Alzheimers Res. Ther.* 6:18. doi: 10.1186/alzrt248
- Oxtoby, N. P., Young, A. L., Cash, D. M., Benzinger, T. L. S., Fagan, A. M., Morris, J. C., et al. (2018). Data-driven models of dominantly-inherited Alzheimer's disease progression. *Brain* 141, 1529–1544. doi: 10.1093/brain/awy050
- Passamonti, L., Vázquez Rodríguez, P., Hong, Y. T., Allinson, K. S. J., Williamson, D., Borchert, R. J., et al. (2016). [18F]AV-1451 positron emission tomography in Alzheimer's disease and progressive supranuclear palsy. *Brain* 140, 781–791.
- Passamonti, L., Vázquez Rodríguez, P. M., Hong, Y. T., Allinson, K. S. J., Bevan-Jones, W. R., Williamson, D., et al. (2018). [11C]PK11195 binding in Alzheimer's disease and progressive supranuclear palsy. *Neurology* 90, e1989–e1996. doi: 10.1212/WNL.0000000000005610
- Su, L., Faluy, Y. O., Hong, Y. T., Fryer, T. D., Mak, E., Gabel, S., et al. (2016). Neuroinflammatory and morphological changes in late-life depression: the NIMROD study. *Br. J. Psychiatry* 209, 525–526. doi: 10.1192/bjp.bp.116.190165
- Villemagne, V. L., Burnham, S., Bourgeat, P., Brown, B., Ellis, K. A., Salvado, O., et al. (2013). Amyloid β deposition, neurodegeneration, and cognitive decline in sporadic Alzheimer's disease: a prospective cohort study. *Lancet Neurol.* 12, 357–367. doi: 10.1016/S1474-4422(13)70044-9
- Wasserman, L. (2000). Bayesian model selection and model averaging. *J. Math. Psychol.* 44, 92–107. doi: 10.1006/jmps.1999.1278
- Wes, P. D., Easton, A., Corradi, J., Barten, D. M., Devidze, N., DeCarr, L. B., et al. (2014). Tau overexpression impacts a neuroinflammation gene expression network perturbed in Alzheimer's disease. *PLoS One* 9:e106050. doi: 10.1371/journal.pone.0106050
- Yang, E., Farnum, M., Lobanov, V., Schultz, T., Verbeeck, R., Raghavan, N., et al. (2011). Quantifying the pathophysiological timeline of Alzheimer's disease. *J. Alzheimers Dis.* 26, 745–753. doi: 10.3233/JAD-2011-110551
- Yoshiyama, Y., Higuchi, M., Zhang, B., Huang, S. M., Iwata, N., Saido, T. C., et al. (2007). Synapse loss and microglial activation precede tangles in a P301S tauopathy mouse model. *Neuron* 53, 337–351. doi: 10.1016/j.neuron.2007.01.010
- Young, A. L., Oxtoby, N. P., Daga, P., Cash, D. M., Ourselin, S., Fox, N. C., et al. (2014). A data driven model of biomarker changes in sporadic Alzheimer's disease. *Brain* 137, 2564–2577. doi: 10.1093/brain/awu176

Conflict of Interest Statement: JO'B has been a consultant for GE Healthcare, TauRx, Axon, Axovant, and Lilly and received honoraria for talks from Piramal.

The remaining 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.

Copyright © 2018 Su, Huang, Wang, Rowe and O'Brien. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Using CT Data to Improve the Quantitative Analysis of ^{18}F -FBB PET Neuroimages

Fermín Segovia^{1*}, Raquel Sánchez-Vañó^{2,3}, Juan M. Górriz^{1,4}, Javier Ramírez^{1,4}, Pablo Sopena-Navales², Nathalie Testart Dardel^{5,6}, Antonio Rodríguez-Fernández^{4,5} and Manuel Gómez-Río^{4,5}

¹ Department of Signal Theory, Networking and Communications, University of Granada, Granada, Spain, ² Department of Nuclear Medicine, "9 de Octubre" Hospital, Valencia, Spain, ³ Clinical Medicine and Public Health Doctoral Program of the University of Granada, Granada, Spain, ⁴ Biosanitary Investigation Institute of Granada, Granada, Spain, ⁵ Department of Nuclear Medicine, "Virgen de las Nieves" University Hospital, Granada, Spain, ⁶ Department of Nuclear Medicine, Lausanne University Hospital (CHUV), Lausanne, Switzerland

OPEN ACCESS

Edited by:

J. Arturo García-Horsman,
University of Helsinki, Finland

Reviewed by:

Felix Carbonell,
Biospective Inc., Canada
Gabriel Gonzalez-Escamilla,
Universitätsmedizin Mainz, Germany

*Correspondence:

Fermín Segovia
fsegovia@ugr.es

Received: 15 December 2017

Accepted: 08 May 2018

Published: 07 June 2018

Citation:

Segovia F, Sánchez-Vañó R, Górriz JM, Ramírez J, Sopena-Navales P, Testart Dardel N, Rodríguez-Fernández A and Gómez-Río M (2018) Using CT Data to Improve the Quantitative Analysis of ^{18}F -FBB PET Neuroimages. *Front. Aging Neurosci.* 10:158. doi: 10.3389/fnagi.2018.00158

^{18}F -FBB PET is a neuroimaging modality that is been increasingly used to assess brain amyloid deposits in potential patients with Alzheimer's disease (AD). In this work, we analyze the usefulness of these data to distinguish between AD and non-AD patients. A dataset with ^{18}F -FBB PET brain images from 94 subjects diagnosed with AD and other disorders was evaluated by means of multiple analyses based on *t*-test, ANOVA, Fisher Discriminant Analysis and Support Vector Machine (SVM) classification. In addition, we propose to calculate amyloid standardized uptake values (SUVs) using only gray-matter voxels, which can be estimated using Computed Tomography (CT) images. This approach allows assessing potential brain amyloid deposits along with the gray matter loss and takes advantage of the structural information provided by most of the scanners used for PET examination, which allow simultaneous PET and CT data acquisition. The results obtained in this work suggest that SUVs calculated according to the proposed method allow AD and non-AD subjects to be more accurately differentiated than using SUVs calculated with standard approaches.

Keywords: quantitative analysis, multivariate analysis, florbetaben, Alzheimer's disease, support vector machine, positron emission tomography

1. INTRODUCTION

Alzheimer's disease (AD) is the most common neurodegenerative disease affecting more than 5 million people in the United States (Alzheimer's Association, 2018) and its prevalence in Europe was estimated at 5.05% (Niu et al., 2017). In addition, the number of AD patients is expected to increase during next decades because of the grow of the older population. Fortunately, the development of new drugs has greatly improved the patient's quality of life, especially when the disease is detected at an early stage. Thus, an early and accurate diagnosis of AD is crucial.

The diagnosis of AD is usually supported by neuroimaging data of different modalities. During the last decade, many research studies have demonstrated that both, structural and molecular imaging, can be successfully used to evaluate patients with AD, including early stages and prodromal AD (Johnson et al., 2012). Structural neuroimaging data such as magnetic resonance imaging (MRI) or computed tomography (CT) allow us to estimate the global cerebral volume,

which was found to significantly correlate with the rate of change in mini-mental state examination scores, evidencing clinical relevance to this marker in the disease progression (Frisoni et al., 2010; Khedher et al., 2015). In addition, MRI and CT data can be used to exclude treatable or reversible causes of dementia (normal-pressure hydrocephalus, subdural hematoma, tumors, etc.).

On the other hand, molecular neuroimages have been widely used in differential diagnosis of dementia. For example, Single Photon Emission Computer Tomography (SPECT) or Positron Emission Tomography (PET) have been demonstrated as valuable tools not only to separate AD patients and controls (Segovia et al., 2010; Rathore et al., 2017) but also to monitor the progression of AD (Hanyu et al., 2010). Probably, the most common molecular neuroimaging modality for AD diagnosis is the well-known ^{18}F -Fludeoxyglucose (FDG) PET. These images allow us to analyze the glucose brain metabolism and that way to estimate the neurodegeneration of certain regions of the brain (Illán et al., 2011; Perani et al., 2014; Cabral et al., 2015).

Conversely to ^{18}F -FDG PET, amyloid imaging focuses on the amyloid beta deposits that characterize AD. During last years, several radiotracers have been proposed to examine these AD hallmarks. The N-methyl- ^{11}C -2-(4'-methylaminophenyl)-6-hydroxybenzothiazole, more commonly referred to as Pittsburgh Compound B (PIB), is an amyloid focused radiotracer traditionally used for this purpose (Klunk et al., 2004). This drug is a radioactive analog of thioflavin T, which binds to amyloid plaques with high affinity, however, its reduced half-life (only 20 min) greatly limits its application (Klunk and Mathis, 2008). Recently, new ^{18}F -labeled tracers with similar efficacy to PIB and longer half-life have been FDA approved: ^{18}F -florbetapir in 2012, ^{18}F -flutemetamol in 2013 and ^{18}F -florbetaben (FBB) in 2014. The validity of these radiotracers is supported by recent studies (Landau et al., 2014; Rice and Bisdas, 2017) that emphasize the added value of these radiotracers in discriminating between AD and non-AD patients (Ceccaldi et al., 2018).

In this work, we analyze ^{18}F -FBB PET data from AD and non-AD patients using univariate and multivariate techniques. In order to improve the diagnosis of AD we propose to include in the analysis the information about gray matter neurodegeneration provided by CT images. This approach takes advantage of the majority of PET images are acquired on scanners that allow simultaneous PET and CT data acquisition. Specifically, we propose to calculate standardized uptake values from ^{18}F -FBB PET data using only voxels belonging to gray matter in CT images. Previous works have followed similar approaches (Villemagne et al., 2015; Rullmann et al., 2016) but in those cases non-gray-matter voxels were discarded only for the reference region and they were determined by means of MRI images. The proposed approach was evaluated using a dataset with ^{18}F -FBB PET and CT scans from 94 subjects acquired during a longitudinal study carried out in two hospitals from the Spanish National Health System. The results suggest that using CT data along with ^{18}F -FBB PET neuroimages improves up to 7% the accuracy of separating AD and non-AD patients, compared with using only PET data.

2. MATERIALS AND METHODS

2.1. Participants

Ninety-four (94) subjects with cognitive impairments were recruited in the Cognitive Behavioral Unit of two different tertiary hospitals: the Virgen de las Nieves hospital in Granada, Spain (72 patients) and the 9 de Octubre hospital in Valencia, Spain (22 patients).

Patients were recruited according to the following clinical criteria: patients with persistent or progressive unexplained MCI Albert et al. (2011); Johnson et al. (2013), defined according to revised Petersen criteria (Winblad et al., 2004); patients fulfilling core clinical criteria for possible AD but an atypical clinical course with no documented progression in the patient's records; patients fulfilling these core clinical criteria but with cerebrovascular comorbidity, concomitant pharmacologic, neurologic, or cognitive components (mixed etiology); and those with a history of progressive dementia and atypically early age at onset (≤ 65 years). All patients fulfilled clinical appropriate use criteria for ^{18}F -FBB PET scan according to international consensus (Johnson et al., 2013). Exclusion criteria were: the presence of a metabolic disorder (hypothyroidism, vitamin B12 or folic acid deficiencies), psychiatric pathology (schizophrenia or depression), MRI-diagnosed cerebrovascular disease (infarction or hemorrhage), neurologic disease affecting gnosis (Parkinsonian syndrome, epilepsy, etc.), pregnancy, glycemia > 160 mg/dL, history of substance abuse, or age < 18 years.

Patients were evaluated using standardized neuropsychological examinations that assessed the orientation, attention, memory, executive function, language, visual and constructive functions and behavior (Carnero Pardo, 2007). In addition, a ^{18}F -FBB PET and a CT scan were acquired for each patient. The imaging protocol in both centers complied with international guidelines (Minoshima et al., 2016). Specific details are given in **Table 1**.

After at least 1 year of follow-up, experienced neurologists established a final diagnosis for each patient on the basis of neuropsychological examinations, the visual assessment of the neuroimaging data and the clinical evolution of the patient. Two subgroups were defined: (i) AD patients and (ii) healthy subjects or patients with diseases other than AD. **Table 2** shows the group distribution and some demographic details of the patients. Note that the second group is very heterogeneous and includes patients with Parkinson's disease, progressive supranuclear palsy and psychiatry disorders among other conditions.

Each patient (or a close relative) gave written informed consent to participate in the study and the protocol was accepted by the Ethics Committee of the "Virgen de las Nieves" hospital (Granada, Spain) and the "9 de Octubre" hospital (Valencia, Spain). All the data were anonymized by the clinicians who acquired them before being considered in this work.

2.2. Data Preprocessing

CT brain images were segmented using the unified segmentation algorithm (Ashburner and Friston, 2005) implemented in Statistical Parametric Mapping (SPM) version 12. This algorithm

TABLE 1 | Protocol details to acquire neuroimaging data.

| | Center A | Center B |
|-----------------------|---|---|
| Camera | GE Discovery STE | Siemens Biograph 16 |
| Patient position | Resting, with closed eyes | Resting, with closed eyes |
| Operation | 3D mode | 3D mode |
| Filtering | Z-Axis standard | Z-Axis standard |
| Dose | 300 MBq | 300 MBq |
| Acq. start | 90 min post injection | 90 min post injection |
| Acq. duration | 20 min | 20 min |
| Matrix size (FBB) | 168 × 168 | 168 × 168 |
| Slice thickness (FBB) | 4.01 mm | 4.06 mm |
| Number of slices | 70 | 70 |
| Voxel size | 16.08 (mm ³) | 16.48 (mm ³) |
| Reconstruction | VUE Point (5 it, 35 sub) | Gaussian + OS-OM (6 it, 21 sub) |
| CT Parameters | Low-dose, 80 mAs, 120 kV | Low-dose, 50 mAs, 120 kV |
| Matrix size (CT) | 512 × 512 | 512 × 512 |
| Slice thickness (CT) | 1 mm | 1 mm |
| Corrections | Scatter; CT attenuation; well counter sensitivity and activity; delayed event subtraction and normalization | Scatter; CT attenuation; slice coincidence location with CT |

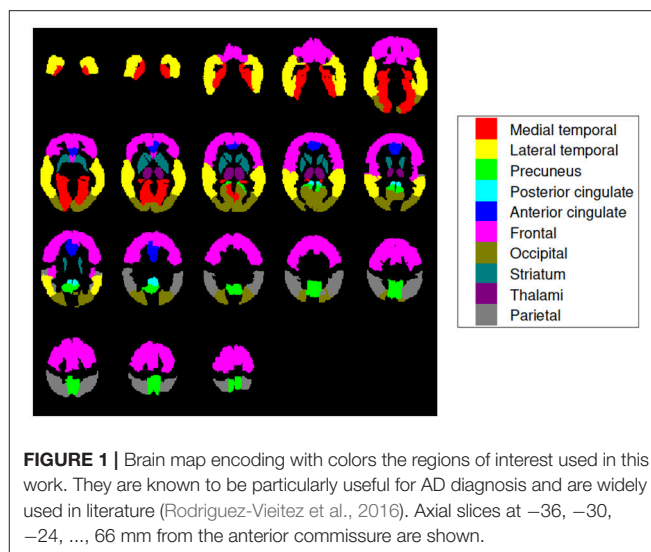
TABLE 2 | Demographic details of the patients considered in this work (μ and σ stand for the average and the standard deviation respectively).

| | # | Sex | | Age | | |
|--------|----|-----|----|-------|----------|-------|
| | | M | F | μ | σ | range |
| AD | 51 | 23 | 28 | 63.43 | 6.32 | 49–74 |
| Non-AD | 43 | 28 | 15 | 62.91 | 8.27 | 42–79 |

allows the separation of gray matter, white matter and cerebrospinal fluid tissues from CT images. The ^{18}F -FBB PET images were also registered to a common space using SPM. This procedure made use of the deformation fields obtained during the segmentation of the CT data in order to achieve a more accurate transformation (Ashburner and Friston, 2007). As a result, we got brain images in Montreal Neuroimaging Institute (MNI) space with $121 \times 141 \times 121$ voxels of $1.5 \times 1.5 \times 1.5$ mm each.

2.3. Regions of Interest

Ten (10) regions of interest (ROIs) were defined to analyze our ^{18}F -FBB PET data: medial temporal, lateral temporal, precuneus, posterior cingulate, anterior cingulate, frontal, occipital, striatum, thalami, and parietal (Rodríguez-Vieitez et al., 2016). Locations and sizes can be seen in **Figure 1**. These regions are frequently associated to AD in the literature and allow comparing our results with the ones obtained by other works (Villemagne et al., 2012; Daerr et al., 2016; Tiepolt et al., 2016; Tuszyński et al., 2016; Bullich et al., 2017). In order to parcel these target regions in our brain images the Automatic Anatomical Labeling (AAL) atlas was used (Tzourio-Mazoyer et al., 2002).

**FIGURE 1** | Brain map encoding with colors the regions of interest used in this work. They are known to be particularly useful for AD diagnosis and are widely used in literature (Rodríguez-Vieitez et al., 2016). Axial slices at -36 , -30 , -24 , ..., 66 mm from the anterior commissure are shown.

2.4. Quantification of ^{18}F -FBB PET Data Using Structural Information

In the clinical practice, neuroimaging data are usually analyzed in terms of standardized uptake values (SUV), which are often given as a ratio of the uptake of a reference region (SUVR). Different regions have been proposed to be used as reference to calculate SUVRs from amyloid PET data (Brendel et al., 2015; Klein et al., 2015a,b; Kimura et al., 2016; Shokouhi et al., 2016). Despite there is no general consensus, the use of the whole cerebellum (Daerr et al., 2016; Bullich et al., 2017) or the cerebellar gray matter (Villemagne et al., 2015) is usually accepted. The SUVR for a given region, k , could be computed as:

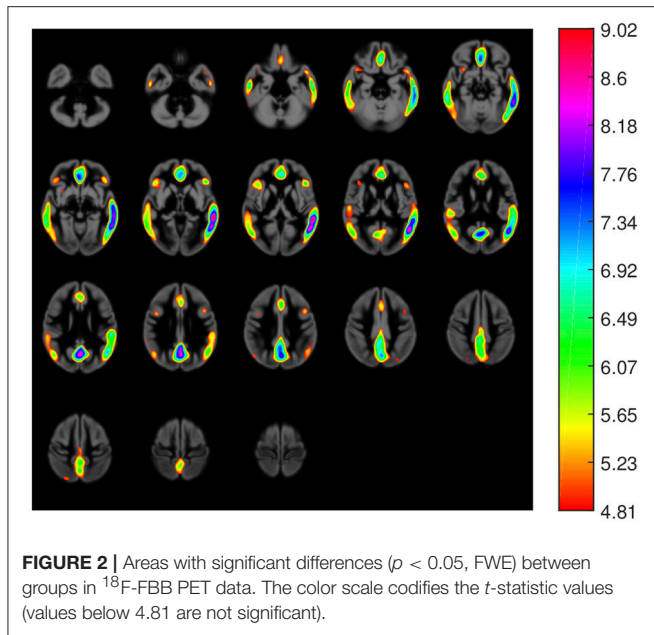
$$\text{SUVR}_k = \frac{N_r \sum_{i=1}^{N_k} x_i}{N_k \sum_{j=1}^{N_r} x_j} \quad (1)$$

where x_i is the intensity of the i -th voxel belonging to region k , with $i \in [1, 2, \dots, N_k]$ and similarly, x_j stands for the intensity of the j -th voxel belonging to a reference region, with $j \in [1, 2, \dots, N_r]$. In this work, we used the whole cerebellum as reference region thus, SUVRs for a given subject were weighted by the mean cerebellar intensity of that subject. This analysis is somewhat similar to the visual examination of the data traditionally performed by experienced clinicians.

Instead of SUVR described by Equation 1, we propose to use a similar measure that also takes into account structural data. Specifically we propose to compute SUVRs using only voxels belonging to gray matter, i.e., those whose position corresponds to gray-matter voxels in CT data. That way we consider not only the amyloid deposits but also the brain injury caused by the disorder.

2.5. Fisher's Discriminant Analysis

The Fisher's discriminant ratio, J , (Theodoridis and Koutroumbas, 2008) is a statistical measure widely used to maximize the differences of means in between two or more



classes respective to the within class variance (Lopez et al., 2009). Mathematically it is defined as:

$$J(\mathbf{w}) = \frac{\mathbf{w}^T S_B \mathbf{w}}{\mathbf{w}^T S_W \mathbf{w}} \quad (2)$$

where \mathbf{w} represent a direction in the data space and S_B and S_W are respectively the “between classes” and the “within classes” scatter matrices. Note that scatter matrices are proportional to covariance matrices and, when only 2 classes are considered, S_B can be expressed as:

$$S_B = (\mu_1 - \mu_2)(\mu_1 - \mu_2)^T \quad (3)$$

where μ_i denotes the mean of the samples belonging to the i -th class. This analysis was not applied to individual voxels (each possible direction in the image space would correspond to a specific voxel position) but to the SUVRs of the ROIs defined in section 2.3. Thus, J was computed as:

$$J_r = \frac{(\mu_1^{(r)} - \mu_2^{(r)})^2}{(\sigma_1^{(r)})^2 + (\sigma_2^{(r)})^2} \quad (4)$$

where $\mu_i^{(r)}$ and $\sigma_i^{(r)}$ are, respectively, the mean and the standard deviation of the SUVR of region r for subjects belonging to the i -th class.

2.6. Support Vector Machine

A binary classification method is a statistical procedure intended to assign a binary label (defining a category or group) to unseen patterns represented by a set of features. To this end, supervised

TABLE 3 | AD target regions and areas with significant differences between AD and non-AD patients in ^{18}F -FBB PET data.

| Region name | Region size | Area with significant diff. |
|---------------------|-------------|-----------------------------|
| Medial temporal | 47171 mm | 432 mm (0.92 %) |
| Lateral temporal | 94907 mm | 36732 mm (38.70 %) |
| Precuneus | 24338 mm | 9693 mm (39.83 %) |
| Posterior cingulate | 2813 mm | 551 mm (19.57 %) |
| Anterior cingulate | 9603 mm | 5346 mm (55.67 %) |
| Frontal | 193730 mm | 10424 mm (5.38 %) |
| Occipital | 50117 mm | 6248 mm (12.47 %) |
| Striatum | 16326 mm | 0 mm (0.00 %) |
| Thalami | 7422 mm | 0 mm (0.00 %) |
| Parietal | 50763 mm | 4715 mm (9.29 %) |

The differences were determined by means of a t -test analysis.

methods build a function $f: \mathbb{R}^D \rightarrow \pm 1$ using a set of known patterns, \mathbf{x}_i and their labels, y_i (training data):

$$(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2), \dots, (\mathbf{x}_N, y_N) \in (\mathbb{R}^k \times \{\pm 1\}) \quad (5)$$

Support Vector Machine (SVM) is a supervised classifier derived from the statistical learning theory (Vapnik, 1998). In SVM the classification function is built using a hyperplane, called maximal margin hyperplane, that has the largest distance to the closest training data pattern of any group:

$$g(\mathbf{x}) = \mathbf{w}^T \mathbf{x} + w_0 = 0, \quad (6)$$

where \mathbf{w} is the weight vector, orthogonal to the decision hyperplane, and w_0 is the threshold. SVM is able to work in combination with kernel approaches when the linear separation of the data is not possible (Müller et al., 2001). Once the hyperplane is computed the classifier assigns a group label to each new pattern according to the side of the hyperplane where it is.

In our experiments the cost parameter, C , was fixed to the commonly accepted value of $C = 1$ and only linear kernels were used. The evaluation of the classification performance was carried out using a 10-fold cross-validation scheme (Varma and Simon, 2006). Given that we have data from 94 subjects, each fold uses 85 samples for training and 9 for test. In the training step of each fold, SUVRs or voxel intensities from each training subject and a binary label determining the group the subject belongs to (AD or non-AD) were used as input data (variables \mathbf{x}_i and y_i in Equation 5). In the test step, the classifier was used to estimate a label for each test subject (represented by its SUVRs or voxel intensities). The estimated labels were then compared with the real ones to assess the classification performance.

2.7. ROC Analysis

In a classification procedure, the trade off between sensitivity and specificity can be analyzed through a Receiver Operating Characteristic (ROC) curve (Brown and Davis, 2006). In these plots, each point represents a sensitivity/specificity pair corresponding to a particular decision threshold. The upper

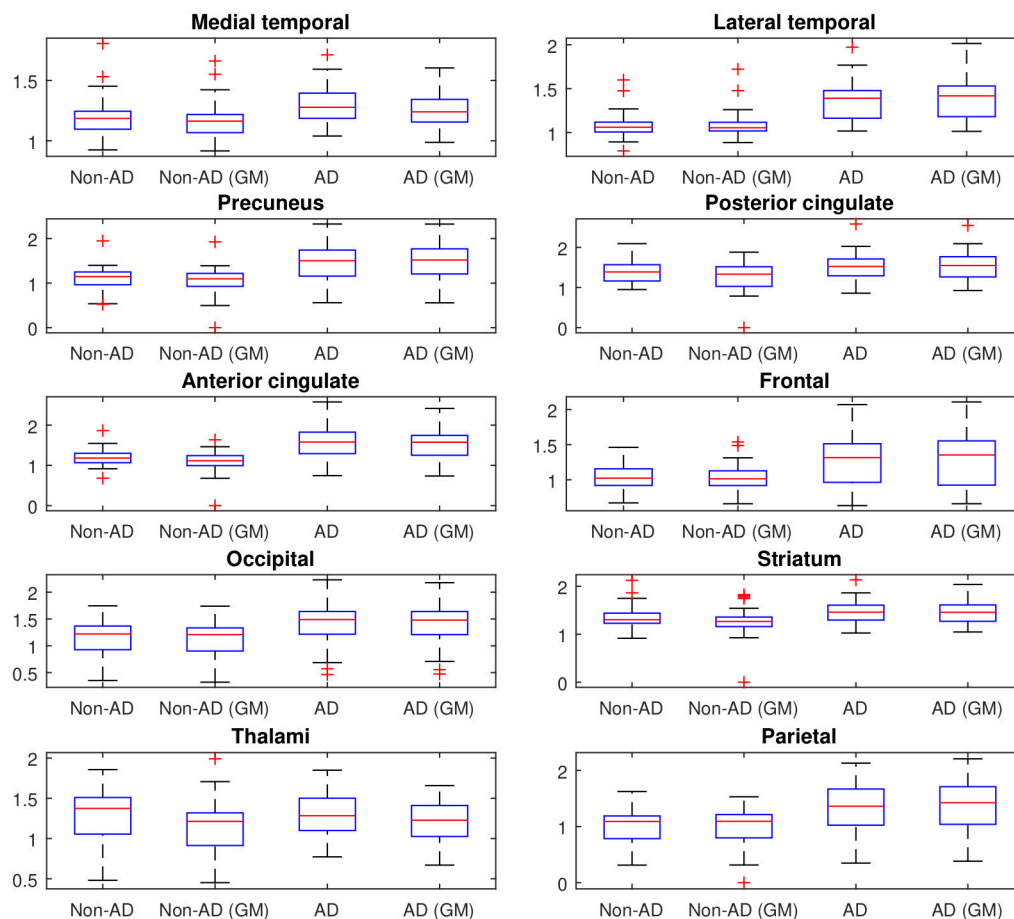


FIGURE 3 | SUVR of the 10 target regions described in section 2.3. The values are grouped into 4 groups according to: (i) the class they belong to (AD or non-AD) and (ii) how they were calculated (using all voxels (classical approach) or using only gray-matter voxels (proposed approach)). On each blue box, the central mark indicates the median, and the bottom and top edges of the box indicate the 25th and 75th percentiles, respectively.

left corner correspond to a sensitivity and specificity of 100%, therefore, the closer the ROC curve is to the upper left corner, the highest accuracy. The area under the curve (AUC) allows measuring how close is the solution to the optimal one and is frequently used as a measure of the classification performance.

3. EXPERIMENT AND RESULTS

First, we carried out a *t*-test analysis on SPM to look for group differences between AD and non-AD subjects for both, ^{18}F -FBB PET and CT data. As suggested in Friston et al. (2006), a smoothed version of the brain images (Gaussian filter of 8 mm FWHM) was used. Results for PET images are shown in **Figure 2**. In this case, we evaluated the hypothesis that data from AD patients have higher intensity than those from non-AD subjects (AD patients are expected to have a greater amyloid-beta concentration). Voxels with significant differences ($p < 0.05$, FWE) between both groups are shown in a specific color which depends on its *t*-statistic. In order to determine if the colored regions match with target regions described in section 2.3, we calculated the percentage of those regions covered by

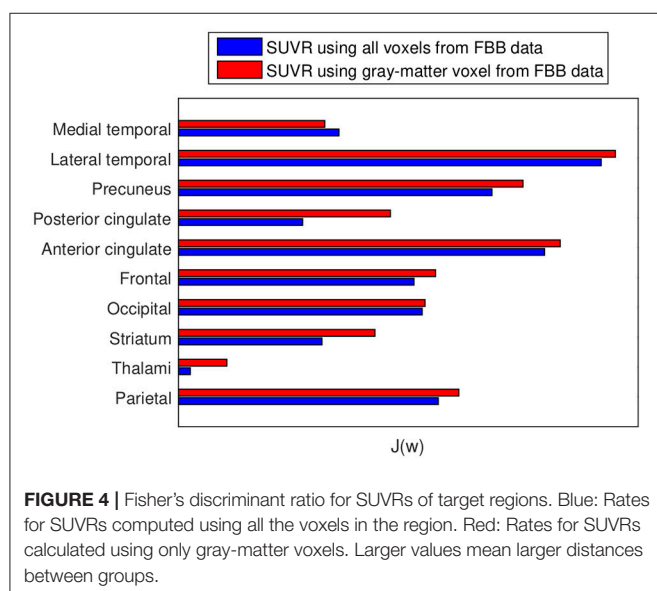
colored voxels. The results are shown in **Table 3**. No significant effects were found for CT data. In this case, only the gray-matter was used and two hypotheses were evaluated: non-AD group has lower intensity than AD group (same as for PET images) and AD group have lower intensity than non-AD group. The latter hypothesis was the most plausible for CT images since one might expect a greater neurodegeneration in AD patients.

Afterwards, the advantages of computing SUVRs from ^{18}F -FBB PET data using only the gray-matter voxels were evaluated. **Figure 3** shows the median SUVR of each target region grouped into four groups according to: i) the class they belong to (AD or non-AD) and ii) how they were calculated: using all voxels (classical approach) or using only gray-matter voxels (proposed approach). The F-statistic (ANOVA) and corresponding *p*-value were computed to determine whether AD and non-AD subjects have different mean on target regions. Results using the classical and the proposed procedure to calculate SUVRs are given in **Table 4**.

The advantages of proposed SUVRs were also assessed by means of the Fisher's discriminant analysis. J_r values (Equation 4)

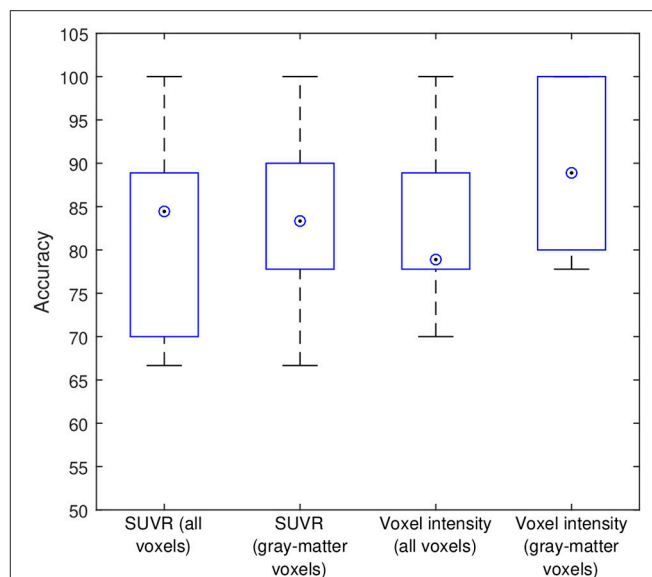
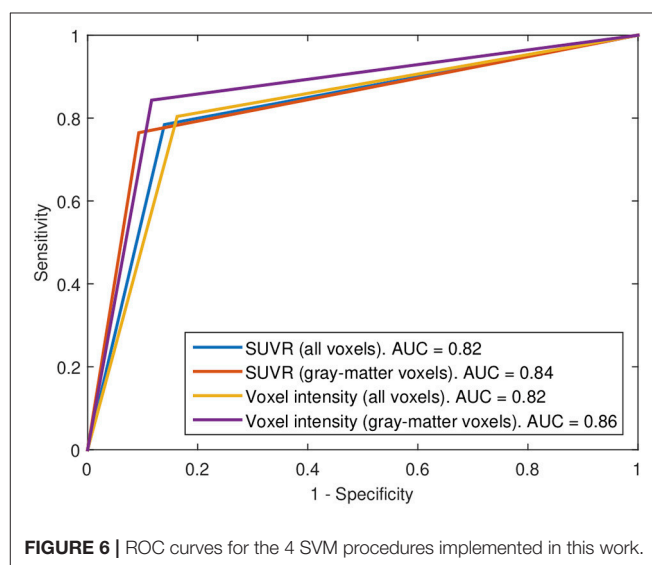
TABLE 4 | F-statistics and corresponding *p*-values to determine whether AD and non-AD patients have different mean on target regions.

| Region | Classical SUVR | | Proposed SUVR | |
|---------------------|----------------|-----------------|---------------|-----------------|
| | F-statistic | <i>p</i> -value | F-statistic | <i>p</i> -value |
| Medial temporal | 7.9922 | 0.0058 | 6.6137 | 0.0117 |
| Lateral temporal | 50.3387 | 0.0000 | 53.6227 | 0.0000 |
| Precuneus | 27.7957 | 0.0000 | 34.3257 | 0.0000 |
| Posterior cingulate | 4.5438 | 0.0357 | 13.6022 | 0.0004 |
| Anterior cingulate | 37.1421 | 0.0000 | 41.8245 | 0.0000 |
| Frontal | 15.2235 | 0.0002 | 18.1024 | 0.0001 |
| Occipital | 17.4945 | 0.0001 | 17.9269 | 0.0001 |
| Striatum | 6.2384 | 0.0143 | 12.1945 | 0.0007 |
| Thalami | 0.0437 | 0.8349 | 0.7259 | 0.3964 |
| Parietal | 19.0551 | 0.0000 | 22.6125 | 0.0000 |

**FIGURE 4** | Fisher's discriminant ratio for SUVRs of target regions. Blue: Rates for SUVRs computed using all the voxels in the region. Red: Rates for SUVRs calculated using only gray-matter voxels. Larger values mean larger distances between groups.

were computed to rate the usefulness of SUVRs of target regions when separating AD and non-AD subjects. **Figure 4** allows us to compare the J_r values computed using all brain voxels with those that considered only gray-matter voxels.

Subsequently, our data were analyzed in terms of their usefulness to separate AD and non-AD patients using SVM classification. Specifically, we estimated the accuracy, sensitivity and specificity of a SVM classifier that separates the groups using ^{18}F -FBB PET data. Two approaches were applied: (i) using SUVR of target regions as feature and (ii) using the intensity of all the voxels in brain images as feature. In both cases we compared the classification results when using or not the CT data to exclude non-gray-matter voxels. For the approach using all the voxels in brain images, the intensity of the voxels was referenced to the mean uptake of the whole cerebellum. This is similar to the intensity normalization performed during the calculation of SUVRs but, in this case, the normalization was individually applied to each voxel. The classification results are shown in **Table 5** and **Figure 5**. The trade off between sensitivity

**FIGURE 5** | Intermediate accuracies obtained in the cross-validation procedure. Blue boxes and circled dots represent accuracies' range and median respectively.**FIGURE 6** | ROC curves for the 4 SVM procedures implemented in this work.

and specificity of the SVM analyses was examined by means of ROC curves. They are shown along with the AUC in **Figure 6**.

The weight map calculated by SVM (parameter w in Equation 6) allows us to examine the importance of each feature in the classification procedure. SVM weights from systems using SUVRs as feature are shown in **Table 6** whereas those calculated by systems using voxel intensities as feature are shown in **Figure 7**. Note that in former systems only 10 regions were used, thus only 10 weights were calculated. Similarly, the systems using voxel intensities as feature computed as many weights as voxels were used.

All the experiments were carried out on Matlab using its statistical toolbox and specific *ad hoc* routines.

4. DISCUSSION

The experiments carried out in this work corroborated that ¹⁸F-FBB PET is an useful neuroimaging modality to assist the diagnosis of AD. Both, univariate and multivariate analyses indicated that these data allow us to separate AD and non-AD subjects with high accuracy. In addition, the regions commonly focused on AD diagnosis show large group differences in ¹⁸F-FBB

TABLE 5 | Classification measures obtained by a SVM classifier when separating AD and non-AD subjects using ¹⁸F-FBB PET data.

| Data used as feature | Accuracy (%) | Sensitivity (%) | Specificity (%) |
|--------------------------------------|--------------|-----------------|-----------------|
| SUVR (all voxels) | 81.91 | 78.43 | 86.05 |
| SUVR (gray-matter voxels) | 82.98 | 76.47 | 90.70 |
| Voxel intensity (all voxels) | 81.91 | 80.39 | 83.72 |
| Voxel intensity (gray-matter voxels) | 86.17 | 84.31 | 88.37 |

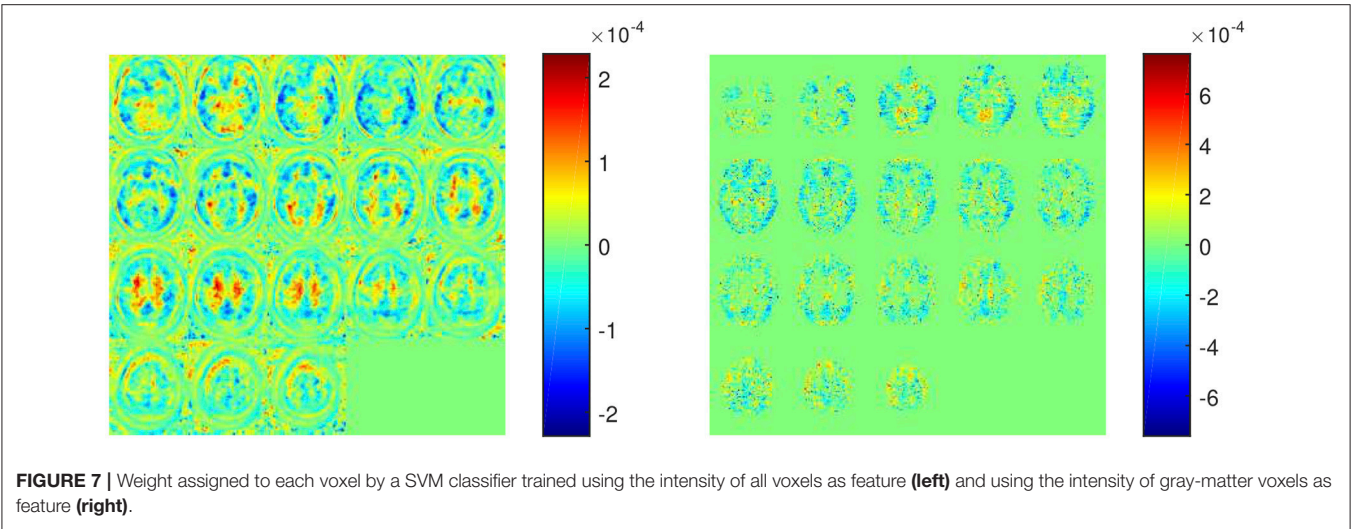
TABLE 6 | Weight assigned to each target region by a SVM classifier that used the SUVRs of those regions as feature.

| ROI | All voxels | Gray-matter voxels |
|---------------------|------------|--------------------|
| Medial temporal | 0.09 | 0.17 |
| Lateral temporal | −1.90 | −2.28 |
| Precuneus | −0.68 | −0.88 |
| Posterior cingulate | 0.37 | 0.50 |
| Anterior cingulate | −0.81 | −0.90 |
| Frontal | 1.10 | 1.45 |
| Occipital | 0.54 | 0.55 |
| Striatum | 0.31 | 0.48 |
| Thalami | 0.29 | −0.16 |
| Parietal | 0.12 | 0.16 |

Two approaches to calculate SUVRs were assessed: using all the voxels in the region (center column) and using only gray-matter voxels in the region (right column).

PET neuroimages. According to the results shown in Table 3, lateral temporal, precuneus, posterior and anterior cingulate have significant differences between groups. Additionally, the former region is the more important one to separate the AD and non-AD subjects as suggested by the results shown in Table 6. The results shown in these tables should be carefully interpreted. Table 3 contains the percentage of each ROI with significant differences ($p < 0.05$, FWE) whereas Table 6 shows the weights assigned by a SVM classifier to those regions when the SUVRs of those regions were used to train the classifier. Thus, frontal is an important region in the separation problem because the classifier assigned it a high weight (relatively high compared with other weights). However, only about 5.38% of voxels in this region (according to the AAL atlas) showed significant differences in the t -test. This suggest that the importance of frontal in the separation problem is not homogeneous throughout the region and some frontal “subregions” are more important than others. It should be noted that frontal was defined as a big region (with a volume of almost 200 cm³ in the AAL atlas), more than 10 times larger than precuneus, a small region with high significance but with not such a high weight. Lateral temporal and anterior cingulate are also two important regions because of their large absolute value in Table 6. SVM weights concern the side of the hyperplane where patterns are placed. Thus, negative weights are associated to regions that characterize non-AD subjects (they “move” patterns toward the non-AD space) whereas positive weights are associated to AD subjects (they “move” patterns toward the AD space) (Caragea et al., 2001). Observe that using only gray-matter voxels made the weights more positive or more negative for all regions except for thalami. This suggests that ¹⁸F-FBB PET data contain no important information to separate the groups in this region. This is consistent with t -test results, which found no significant differences in thalami.

The t -test analysis has drawn two clear conclusions: (i) there are significant group differences in ¹⁸F-FBB PET neuroimages and, (ii) there are no significant group differences in CT data. The latter conclusion can be explained by the composition of



the non-AD group, which contains a large proportion of subjects with other diseases, including parkinsonian disorders, that could have structural changes similar to AD. The lack of significant group differences in CT data may also be due to the neuroimaging modality itself (Gado et al., 1983). Although a number of studies (Grundman et al., 2002; Rathore et al., 2017) have reported volumetric differences between AD and non-AD patients in MRI data, the use of CT neuroimages to this purpose have been poorly studied.

In this work we propose to use SUVRs from ^{18}F -FBB PET neuroimages that also consider the gray matter neurodegeneration in order to improve the diagnosis of AD. In most cases, this information can be extracted from CT data in a inexpensive and efficient way, since most of the scanners used for PET are combined PET/CT devices. Specifically, we propose to discard those voxels from ^{18}F -FBB PET images not belonging to gray matter in CT images and therefore, calculate SUVRs using only the gray-matter voxels. The idea of discarding non-gray-matter voxels was used in previous works (Villemagne et al., 2015; Rullmann et al., 2016) to calculate the SUV of the reference region or to perform intermediate corrections. Here, we propose to apply it to the SUV calculation of all the regions and, to this end, we propose to use CT data due to its greater availability. This way to compute SUVs is similar to the one used in Gonzalez-Escamilla et al. (2017) but we used CT instead MRI images. The results obtained in this work suggest that the proposed approach allows separating AD and non-AD patients more accurately than using standard methods for SUVR calculation. As shown in **Figure 4**, for 9 out of 10 ROIs the computation of the SUVR that considered only the gray matter separated the patient groups more than the SUVR computed using standard methods. These results were corroborated by ANOVA and SVM analyses. **Tables 4, 5** show that mean differences between groups are greater (higher F-statistics and lower *p*-values) and accuracy rates in SVM classification are larger when SUVRs were computed using only gray-matter voxels.

SVM classification performed an accurate separation (accuracy above 80% for the 4 studied feature sets) of the groups, which is particularly important if we take into account that the separation of AD patients from other neurological disorders is more difficult than distinguishing between AD patients and healthy subjects (as mentioned before non-AD group contains a large number of patients with other disorders). Although the heterogeneity of non-AD group could be seen as a limitation of our study, this approach is, in our opinion, more interesting because it is very similar to the clinical problem where clinicians usually take care of non-healthy subjects and should differentiate between AD and other disorders. The obtained accuracy rates

suggest that ^{18}F -FBB PET data contain useful biomarkers to develop computer-aided diagnosis systems for AD. Anyway, the analysis of the reported accuracy rates should consider potential labeling errors inherent in all diagnostics.

The proposed approach to calculate SUVRs must not be confused with Partial Volume Effect correction (PVEc) methods (Erlandsson et al., 2012; Matsubara et al., 2016; Rullmann et al., 2016). In fact, the application of those corrections are compatible with the way to calculate SUVRs that we are proposing. In this work, we decided not using PVEc methods due to: (i) presently, they are not routinely applied, neither in clinical nor in research settings and (ii) these techniques depend on a range of model assumptions and may result on noise amplification (Erlandsson et al., 2012; Greve et al., 2016; Gonzalez-Escamilla et al., 2017).

5. CONCLUSIONS

In this work we have proposed to compute SUVRs from amyloid-PET imaging considering also structural data. Specifically, we proposed to use only gray-matter voxels, estimated through CT images, to calculate SUVRs. In order to evaluate the proposed approach, different experiments based on *t*-test, ANOVA, FDR and SVM were carried out. A dataset with ^{18}F -FBB PET and CT brain images from 94 subjects diagnosed with AD and other disorders was used for evaluation purposes. The results of those experiments suggest that the proposed method to calculate SUVRs allows separating AD and non-AD subjects more accurately than SUVRs calculated by standard methods. Additionally, the results obtained in this work corroborated that ^{18}F -FBB PET data are good biomarkers to estimated brain amyloid deposits and are useful to diagnose AD.

AUTHOR CONTRIBUTIONS

RS-V, PS-N, NT-D, AR-F, and MG-R: Data collection; FS, RS-V, JG, JR, PS-N, and MG-R: Conception and design of the work; FS, JG, and JR: Data analysis; FS, RS-V, JG, JR, PS-N, and MG-R: Result interpretation; FS: Drafting the article; FS, RS-V, JG, JR, PS-N, and MG-R: Critical revision of the article.

ACKNOWLEDGMENTS

This work was supported by the MINECO/FEDER under the TEC2012-34306 and TEC2015-64718-R projects and the Ministry of Economy, Innovation, Science and Employment of the Junta de Andalucía under the Excellence Project P11-TIC-7103. The work was also supported by the Vicerectorate of Research and Knowledge Transfer of the University of Granada.

REFERENCES

- Albert, M. S., DeKosky, S. T., Dickson, D., Dubois, B., Feldman, H. H., Fox, N. C., (2011). The diagnosis of mild cognitive impairment due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement.* 7, 270–279. doi: 10.1016/j.jalz.2011.03.008
- Alzheimer's Association (2018). 2018 Alzheimer's disease facts and figures. *Alzheimers Dement.* 14, 367–429. doi: 10.1016/j.jalz.2018.02.001
- Ashburner, J., and Friston, K. (2007). "Non-linear Registration," in *Statistical Parametric Mapping*, eds K. Friston, J. Ashburner, S. Kiebel, T. Nichols, and W. Penny (London: Academic Press), 63–80.
- Ashburner, J., and Friston, K. J. (2005). Unified segmentation. *NeuroImage* 26, 839–851. doi: 10.1016/j.neuroimage.2005.02.018

- Brendel, M., Högenauer, M., Delker, A., Sauerbeck, J., Bartenstein, P., Seibyl, J., et al. (2015). Improved longitudinal [18F]-AV45 amyloid PET by white matter reference and VOI-based partial volume effect correction. *NeuroImage* 108, 450–459. doi: 10.1016/j.neuroimage.2014.11.055
- Brown, C. D., and Davis, H. T. (2006). Receiver operating characteristics curves and related decision measures: A tutorial. *Chemomet. Intell. Lab. Syst.* 80, 24–38. doi: 10.1016/j.chemolab.2005.05.004
- Bullich, S., Villemagne, V. L., Catafau, A. M., Jovalekic, A., Koglin, N., Rowe, C. C., et al. (2017). Optimal reference region to measure longitudinal amyloid-beta change with 18F-Florbetaben PET. *J. Nuclear Med.* 58, 1300–1306. doi: 10.2967/jnumed.116.187351
- Cabral, C., Morgado, P. M., Campos Costa, D., and Silveira, M. (2015). Predicting conversion from MCI to AD with FDG-PET brain images at different prodromal stages. *Comput. Biol. Med.* 58, 101–109. doi: 10.1016/j.combiomed.2015.01.003
- Caragea, D., Cook, D., and Honavar, V. G. (2001). “Gaining insights into support vector machine pattern classifiers using projection-based tour methods,” in *Proceedings of the Seventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '01, (New York, NY: ACM), 251–256.
- Carnero Pardo, C. (2007). Utilidad diagnóstica del Test de las Fotos (Fototest) en el deterioro cognitivo y demencia. *Neurología*. 22, 860–869.
- Ceccaldi, M., Jonveaux, T., Verger, A., Krolak-Salmon, P., Houzard, C., Godefroy, O., et al. (2018). Added value of 18F-florbetaben amyloid PET in the diagnostic workup of most complex patients with dementia in France: A naturalistic study. *Alzheimer's Dement.* 14, 293–305. doi: 10.1016/j.jalz.2017.09.009
- Daerr, S., Brendel, M., Zach, C., Mille, E., Schilling, D., Zacherl, M. J., et al. (2016). Evaluation of early-phase [18F]-florbetaben PET acquisition in clinical routine cases. *NeuroImage* 14, 77–86. doi: 10.1016/j.nicl.2016.10.005
- Erlundsson, K., Buvat, I., Pretorius, P. H., Thomas, B. A., and Hutton, B. F. (2012). A review of partial volume correction techniques for emission tomography and their applications in neurology, cardiology and oncology. *Phys. Med. Biol.* 57:R119. doi: 10.1088/0031-9155/57/21/R119
- Frisoni, G. B., Fox, N. C., Jack, C. R., Scheltens, P., and Thompson, P. M. (2010). The clinical use of structural MRI in Alzheimer disease. *Nat. Rev. Neurol.* 6, 67–77. doi: 10.1038/nrneurol.2009.215
- Friston, K. J., Ashburner, J. T., Kiebel, S. J., Nichols, T. E., and Penny, W. D. (2006). *Statistical Parametric Mapping: The Analysis of Functional Brain Images*. Amsterdam; Boston: Academic Press.
- Gado, M., Patel, J., Hughes, C. P., Danziger, W., and Berg, L. (1983). Brain atrophy in dementia judged by CT scan ranking. *AJNR* 4, 499–500.
- Gonzalez-Escamilla, G., Lange, C., Teipel, S., Buchert, R., and Grothe, M. J. (2017). PETPVE12: An SPM toolbox for Partial Volume Effects correction in brain PET – Application to amyloid imaging with AV45-PET. *NeuroImage* 147, 669–677. doi: 10.1016/j.neuroimage.2016.12.077
- Greve, D. N., Salat, D. H., Bowen, S. L., Izquierdo-Garcia, D., Schultz, A. P., Catana, C., et al. (2016). Different partial volume correction methods lead to different conclusions: An 18F-FDG-PET study of aging. *NeuroImage* 132, 334–343. doi: 10.1016/j.neuroimage.2016.02.042
- Grundman, M., Sencakova, D., Jack, C. R., Petersen, R. C., Kim, H. T., Schultz, A., et al. (2002). Brain MRI hippocampal volume and prediction of clinical status in a mild cognitive impairment trial. *J. Mol. Neurosci.* 19, 23–27. doi: 10.1007/s12031-002-0006-6
- Hanyu, H., Sato, T., Hirao, K., Kanetaka, H., Iwamoto, T., and Koizumi, K. (2010). The progression of cognitive deterioration and regional cerebral blood flow patterns in Alzheimer's disease: A longitudinal SPECT study. *J. Neurol. Sci.* 290, 96–101. doi: 10.1016/j.jns.2009.10.022
- Illán, I. A., Górriz, J. M., Ramírez, J., Salas-Gonzalez, D., López, M. M., Segovia, F., et al. (2011). 18F-FDG PET imaging analysis for computer aided Alzheimer's diagnosis. *Inf. Sci.* 181, 903–916. doi: 10.1016/j.ins.2010.10.027
- Johnson, K. A., Fox, N. C., Sperling, R. A., and Klunk, W. E. (2012). Brain Imaging in Alzheimer Disease. *Cold Spring Harb. Perspect. Med.* 2:a006213. doi: 10.1101/cshperspect.a006213
- Johnson, K. A., Minoshima, S., Bohnen, N. I., Donohoe, K. J., Foster, N. L., Herscovitch, P., et al. (2013). Update on appropriate use criteria for amyloid PET imaging: dementia experts, mild cognitive impairment, and education. *J. Nucl. Med.* 54, 1011–1013. doi: 10.2967/jnumed.113.127068
- Khedher, L., Ramírez, J., Górriz, J. M., Brahim, A., and Segovia, F. (2015). Early diagnosis of Alzheimer's disease based on partial least squares, principal component analysis and support vector machine using segmented MRI images. *Neurocomputing* 151 (pt 1), 139–150. doi: 10.1016/j.neucom.2014.09.072
- Kimura, Y., Yamada, T., Hosokawa, C., Okada, S., Nagaoka, T., and Ishii, K. (2016). Delineation algorithm on reference region for amyloid imaging using a time history of radioactivity. *J. Nucl. Med.* 57(suppl. 2), 311–311.
- Klein, G., Sampat, M., Staewen, D., Scott, D., and Suhy, J. (2015a). Comparative assessment of SUVR methods and reference regions in amyloid PET studies. *Alzheimer's Dement.* 11:P350. doi: 10.1016/j.jalz.2015.06.231
- Klein, G., Sampat, M., Staewen, D., Scott, D., and Suhy, J. (2015b). Comparison of SUVR Methods and Reference Regions in Amyloid PET. *J. Nuclear Med.* 56(suppl.3), 1741.
- Klunk, W. E., Engler, H., Nordberg, A., Wang, Y., Blomqvist, G., Holt, D. P., et al. (2004). Imaging brain amyloid in Alzheimer's disease with Pittsburgh Compound-B. *Ann. Neurol.* 55, 306–319. doi: 10.1002/ana.20009
- Klunk, W. E., and Mathis, C. A. (2008). The future of amyloid-beta imaging: a tale of radionuclides and tracer proliferation. *Curr. Opin. Neurol.* 21, 683–687. doi: 10.1097/WCO.0b013e3283168e1a
- Landau, S. M., Thomas, B. A., Thurfjell, L., Schmidt, M., Margolin, R., Mintun, M., et al. (2014). Amyloid PET imaging in Alzheimer's disease: A comparison of three radiotracers. *Eur. J. Nucl. Med. Mol. Imaging* 41, 1398–1407. doi: 10.1007/s00259-014-2753-3
- Lopez, M., Ramirez, J., Gorriz, J., Salas-Gonzalez, D., Alvarez, I., Segovia, F., et al. (2009). Automatic tool for alzheimer's disease diagnosis using PCA and bayesian classification rules. *Electron. Lett.* 45, 389–391. doi: 10.1049/el.2009.0176
- Matsubara, K., Ibaraki, M., Shimada, H., Ikoma, Y., Suhara, T., Kinoshita, T., et al. (2016). Impact of spillover from white matter by partial volume effect on quantification of amyloid deposition with [11C]PiB PET. *NeuroImage* 143, 316–324. doi: 10.1016/j.neuroimage.2016.09.028
- Minoshima, S., Drzezga, A. E., Barthel, H., Bohnen, N., Djekidel, M., Lewis, D. H., et al. (2016). SNMMI Procedure Standard/EANM Practice Guideline for Amyloid PET Imaging of the Brain 1.0. *J. Nucl. Med.* 57, 1316–1322. doi: 10.2967/jnumed.116.174615
- Müller, K., Mika, S., Rätsch, G., Tsuda, K., and Schölkopf, B. (2001). An introduction to kernel-based learning algorithms. *IEEE Trans. Neural Netw.* 12, 181–201. doi: 10.1109/72.914517
- Niu, H., Álvarez-Álvarez, I., Guillén-Grima, F., and Aguina-Ontoso, I. (2017). Prevalence and incidence of Alzheimer's disease in Europe: A meta-analysis. *Neurología* 32, 523–532. doi: 10.1016/j.nrl.2016.02.016
- Perani, D., Della Rosa, P. A., Cerami, C., Gallivanone, F., Fallanca, F., Vanoli, E. G., et al. (2014). Validation of an optimized SPM procedure for FDG-PET in dementia diagnosis in a clinical setting. *NeuroImage* 6, 445–454. doi: 10.1016/j.nicl.2014.10.009
- Rathore, S., Habes, M., Ifthikhar, M. A., Shacklett, A., and Davatzikos, C. (2017). A review on neuroimaging-based classification studies and associated feature extraction methods for Alzheimer's disease and its prodromal stages. *NeuroImage* 155 (Suppl. C), 530–548. doi: 10.1016/j.neuroimage.2017.03.057
- Rice, L., and Bidas, S. (2017). The diagnostic value of FDG and amyloid PET in Alzheimer's disease—A systematic review. *Eur. J. Radiol.* 94, 16–24. doi: 10.1016/j.ejrad.2017.07.014
- Rodriguez-Vieitez, E., Carter, S. F., Chiotis, K., Saint-Aubert, L., Leuzy, A., Schöll, M., et al. (2016). Comparison of early-phase 11C-Deuterium-l-Deprenyl and 11C-Pittsburgh Compound B PET for Assessing Brain Perfusion in Alzheimer Disease. *J. Nucl. Med.* 57, 1071–1077. doi: 10.2967/jnumed.115.168732
- Rullmann, M., Dukart, J., Hoffmann, K.-T., Luthardt, J., Tiepolt, S., Patt, M., et al. (2016). Partial-volume effect correction improves quantitative analysis of 18F-Florbetaben β -Amyloid PET Scans. *J. Nucl. Med.* 57, 198–203. doi: 10.2967/jnumed.115.161893
- Segovia, F., Górriz, J. M., Ramírez, J., Salas-González, D., Álvarez, I., López, M., et al. (2010). Classification of functional brain images using a GMM-based multi-variate approach. *Neurosci. Lett.* 474, 58–62. doi: 10.1016/j.neulet.2010.03.010
- Shokouhi, S., McKay, J. W., Baker, S. L., Kang, H., Brill, A. B., Gwirtsman, H. E., et al. (2016). Reference tissue normalization in longitudinal 18F-florbetapir positron emission tomography of late mild cognitive impairment. *Alzheimer's Res. Ther.* 8:2. doi: 10.1186/s13195-016-0172-3

- Theodoridis, S., and Koutroumbas, K. (2008). *Pattern Recognition, Fourth Edition*. Amsterdam: Academic Press.
- Tiepol, S., Hesse, S., Patt, M., Luthardt, J., Schroeter, M. L., Hoffmann, K.-T., et al. (2016). Early 18F-florbetaben and 11C-PiB PET images are a surrogate biomarker of neuronal injury in Alzheimer's disease. *Eur. J. Nuclear Med. Mol. Imaging* 43, 1700–1709. doi: 10.1007/s00259-016-3353-1
- Tuszynski, T., Rullmann, M., Luthardt, J., Butzke, D., Tiepol, S., Gertz, H.-J., et al. (2016). Evaluation of software tools for automated identification of neuroanatomical structures in quantitative β -amyloid PET imaging to diagnose Alzheimer's disease. *Eur. J. Nucl. Med. Mol. Imaging* 43, 1077–1087. doi: 10.1007/s00259-015-3300-6
- Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., et al. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *NeuroImage* 15, 273–289. doi: 10.1006/nimg.2001.0978
- Vapnik, V. N. (1998). *Statistical Learning Theory*. New York, NY: Wiley-Interscience.
- Varma, S., and Simon, R. (2006). Bias in error estimation when using cross-validation for model selection. *BMC Bioinformatics* 7:91. doi: 10.1186/1471-2105-7-91
- Villemagne, V. L., Bourgeat, P., Doré, V., Macaulay, L., Williams, R., Ames, D., et al. (2015). Amyloid imaging in therapeutic trials: The quest for the optimal reference region. *Alzheimer's Dement.* 11, P21–P22. doi: 10.1016/j.jalz.2015.06.036
- Villemagne, V. L., Mulligan, R. S., Pejoska, S., Ong, K., Jones, G., O'Keefe, G., et al. (2012). Comparison of 11C-PiB and 18F-florbetaben for A-Beta imaging in ageing and Alzheimer's disease. *Eur. J. Nuclear Med. Mol. Imaging* 39, 983–989. doi: 10.1007/s00259-012-2088-x
- Winblad, B., Palmer, K., Kivipelto, M., Jelic, V., Fratiglioni, L., Wahlund, L.-O., et al. (2004). Mild cognitive impairment – beyond controversies, towards a consensus: report of the International Working Group on Mild Cognitive Impairment. *J. Int. Med.* 256, 240–246. doi: 10.1111/j.1365-2796.2004.01380.x
- 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.

Copyright © 2018 Segovia, Sánchez-Vañó, Górriz, Ramírez, Sopena-Novales, Testart Dardel, Rodríguez-Fernández and Gómez-Río. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



MRI Characterizes the Progressive Course of AD and Predicts Conversion to Alzheimer's Dementia 24 Months Before Probable Diagnosis

Christian Salvatore¹, Antonio Cerasa² and Isabella Castiglioni^{1*}
for the Alzheimer's Disease Neuroimaging Initiative[†]

¹ Institute of Molecular Bioimaging and Physiology, National Research Council (IBFM-CNR), Milan, Italy, ² Institute of Molecular Bioimaging and Physiology, National Research Council (IBFM-CNR), Catanzaro, Italy

OPEN ACCESS

Edited by:

Juan Manuel Gorriz,
Universidad de Granada, Spain

Reviewed by:

Li Su,
University of Cambridge,
United Kingdom
Guido Gainotti,
Università Cattolica del Sacro Cuore,
Italy

*Correspondence:

Isabella Castiglioni
isabella.castiglioni@ibfm.cnr.it

[†]Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report.

A complete listing of ADNI investigators can be found at: http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf

Received: 15 December 2017

Accepted: 23 April 2018

Published: 24 May 2018

Citation:

Salvatore C, Cerasa A and Castiglioni I (2018) MRI Characterizes the Progressive Course of AD and Predicts Conversion to Alzheimer's Dementia 24 Months Before Probable Diagnosis. *Front. Aging Neurosci.* 10:135. doi: 10.3389/fnagi.2018.00135

There is no disease-modifying treatment currently available for AD, one of the more impacting neurodegenerative diseases affecting more than 47.5 million people worldwide. The definition of new approaches for the design of proper clinical trials is highly demanded in order to achieve non-confounding results and assess more effective treatment. In this study, a cohort of 200 subjects was obtained from the Alzheimer's Disease Neuroimaging Initiative. Subjects were followed-up for 24 months, and classified as AD (50), progressive-MCI to AD (50), stable-MCI (50), and cognitively normal (50). Structural T1-weighted MRI brain studies and neuropsychological measures of these subjects were used to train and optimize an artificial-intelligence classifier to distinguish mild-AD patients who need treatment (AD + pMCI) from subjects who do not need treatment (sMCI + CN). The classifier was able to distinguish between the two groups 24 months before AD definite diagnosis using a combination of MRI brain studies and specific neuropsychological measures, with 85% accuracy, 83% sensitivity, and 87% specificity. The combined-approach model outperformed the classification using MRI data alone (72% classification accuracy, 69% sensitivity, and 75% specificity). The patterns of morphological abnormalities localized in the temporal pole and medial-temporal cortex might be considered as biomarkers of clinical progression and evolution. These regions can be already observed 24 months before AD definite diagnosis. The best neuropsychological predictors mainly included measures of functional abilities, memory and learning, working memory, language, visuoconstructional reasoning, and complex attention, with a particular focus on some of the sub-scores of the FAQ and AVLT tests.

Keywords: artificial intelligence, Alzheimer's disease, clinical trials, magnetic resonance imaging, neuropsychological tests, biomarkers, predictors

INTRODUCTION

According to the World Health Organization, there were 47.5 million people worldwide with dementia in 2015, with 7.7 million new cases each year. The total number of people with dementia is projected to reach 75.6 millions in 2030 and almost triple by 2050 to 135.5 millions (Dementia Statistics, 2015; World Alzheimer Report, 2015; Khan et al., 2017). The most frequent dementia

form is Alzheimer's Disease (AD) (approximately 70%), whose impact on the society in terms of costs as well as quality of life of patients and families is impressive (Khan et al., 2017). There is no AD-modifying treatment available to date, and one third of the population will die with dementia if something does not change in the approach of screening, diagnosis, prognosis and treatment, including more proper design of clinical trials.

Currently, there are indeed more than 500 open clinical studies on AD, according to *ClinicalTrials.gov*. Many other clinical trials have been closed in the past years, few achieved phase III and no one demonstrated a proper success rate. Most of the past clinical trials enrolled people with advanced AD, and clinicians recommended to treat patients at an earlier stage for more effective results. Thus, current clinical trials try to enroll subjects at an early phase of the disease: inclusion criteria are now based on the selection of this specific patient group.

The patient's self-reported experiences and the observed cognitive, functional and behavioral symptomatology due to AD over the longitudinal course of the illness are the current basis for the clinical diagnosis of AD. However, they are insufficient for detecting early AD subjects, considering also that only 33% of subjects with mild cognitive impairment (MCI) progress to AD (Mitchell and Shiri-Feshki, 2009). Furthermore, no standards have been defined on the best neuropsychological outcomes to be measured for this purpose.

For these reasons, clinical trials based only on neuropsychological assessment risk (1) including subjects with early dementia forms that are not caused by AD and (2) lasting several years prior to be completed, when most of the enrolled subjects have clearly progressed to AD. This leads to confounding clinical-trial designs, and cause treatments to be administered on patients who are not really affected by AD.

In 2011, after many scientific evidences, medical-imaging studies were included in the revised diagnostic criteria for AD in order to detect objective signs of disease in the subjects' brain. Being positive to Positron Emission Tomography (PET) with A β - or tau-specific radiotracers is used as an inclusion criterion in most recent clinical trials, with the aim of measuring the presence of brain β -amyloid plaques or tau deposition, the recognized cause of AD pathogenesis. However, these PET studies are expensive, invasive and difficult to be implemented for technical and authorization problems, in particular in non-western countries. Moreover, lack of success in clinical trials of candidate drugs targeting amyloid or tau proteins has led to target alternative mechanisms (e.g., Khan et al., 2017).

Magnetic Resonance Imaging (MRI) is a less expensive technique than PET, non-invasive and more common in both western and non-western regions, and already recommended to detect AD neuronal degeneration and to monitor AD progression in clinical trials (Sperling et al., 2011). However, radiologists are not always able to detect -by visual inspection- the presence of subtle cerebral signs of neurodegeneration in MCI subjects, and even when this is possible, they are not able to predict if a subject will progress or not to AD.

Artificial-intelligence (AI) technology is emerging as an effective tool for automatic, objective and more sensitive

assessment of imaging studies. Specifically, machine-learning (ML) and pattern-recognition techniques have captured the attention of the neuroimaging community as they have been proven able to discover previously unknown patterns in imaging data (Bishop, 2006; Wernick et al., 2010). In other words, these algorithms are able to (1) extract information from imaging data without *a priori* knowledge of where it may be encoded in the images, and (2) combine the information encoded in multiple inter- and intra-domain variables. This information can then be used to design multivariate mathematical models able to automatically predict the diagnostic class of a subject. This characteristic may be of particular usefulness in the context of early diagnosis, when pathological signs are not yet evident by visual inspection (Salvatore et al., 2015a). In the last years, different ML approaches have been applied to the automatic diagnosis and prognosis of AD by means of cerebral MRI studies, showing good performance even at an early stage of the disease (e.g., Cuingnet et al., 2011; Moradi et al., 2015; Salvatore et al., 2015b; Nanni et al., 2016). Furthermore, good results have been obtained to translate the hidden image features used by ML in performing subject classification, which are often typically complex features, counter-intuitive and not meaningful *per se* to clinicians (Haufe et al., 2014; Salvatore et al., 2015b; Huys et al., 2016). Thus, results of ML classification by means of MRI brain images can be more easily interpreted by clinicians and associated to AD pathogenesis.

The aim of this study is to refine the application of ML systems for the characterization of the progressive course of AD and to predict the conversion of MCI to AD, trying to establish how long before it would be possible to predict the diagnosis of probable AD. Application of this approach to longitudinal datasets would enable us to focus on the prognosis rather than the diagnosis and to identify cost-effective biomarkers, which may be targeted for prevention/intervention programs.

MATERIALS AND METHODS

Participants

Data used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database¹. The ADNI was launched in 2003 by the National Institute on Aging (NIA), the National Institute of Biomedical Imaging and Bioengineering (NIBIB), and the Food and Drug Administration (FDA), as a 5-year public private partnership, led by the principal investigator, Michael W. Weiner, MD. The primary goal of ADNI was to test whether serial magnetic resonance imaging (MRI), positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessments subjected to participants could be combined to measure the progression of MCI and early Alzheimer's disease (AD) – see www.adni-info.org.

As specified in the ADNI protocol², each participant was willing, spoke either English or Spanish, was able to perform all

¹adni.loni.usc.edu

²<http://www.adni-info.org/Scientists/ADNISTudyProcedures.html>

test procedures described in the protocol and had a study partner able to provide an independent evaluation of functioning.

Inclusion criteria for cognitively normal (CN) subjects were: Mini Mental State Examination (MMSE) (Folstein et al., 1975) scores between 24 and 30, Clinical Dementia Rating (CDR) of zero (Morris, 1993), and absence of depression, MCI and dementia. Inclusion criteria for MCI were: MMSE scores between 24 and 30, CDR of 0.5, objective memory loss measured by education-adjusted scores on the Logical Memory II subtest of the Wechsler Memory Scale (Wechsler, 1987), absence of significant levels of impairment in other cognitive domains, and absence of dementia. Inclusion criteria for AD were: MMSE scores between 20 and 26, CDR of 0.5 or 1.0, and criteria for probable AD as defined by the National Institute of Neurological and Communicative Disorders and Stroke (NINCDS) e by the Alzheimer's Disease and Related Disorders Association (ADDA) (McKhann et al., 1984; Dubois et al., 2007).

Serial MRI studies were performed to participants from baseline, covering a follow-up period of several years. Each participant was diagnosed at each time point of serial MRI studies.

In the present work, a total of 200 subjects were retrieved from the ADNI database, consisting into 50 subjects with a stable diagnosis of CN state over the 24 months of follow up, 50 subjects with a stable diagnosis of MCI (sMCI), 50 subjects with a stable diagnosis of AD, and 50 subjects with an initial diagnosis of MCI who showed a progression to AD (pMCI).

Two age- and sex-matched groups of subjects were created by grouping, separately, AD with pMCI (100 subjects) and CN with sMCI (100 subjects).

These subjects had all three serial MRI studies at three time points after the baseline: 6, 12, and 24 months.

The 24-months point was chosen as the time-zero point for a stable diagnosis. As a consequence, the three previous time points were reconsidered (and renamed) as *24 months before stable diagnosis*, *18 months before stable diagnosis*, and *12 months before stable diagnosis*.

Demographic and clinical characteristics of the groups of ADNI subjects considered in this study are shown in **Table 1**. ADNI Subject IDs as well as Image Data IDs can be found at the following online repository: <https://github.com/christiansalvatore/Salvatore-200Longitudinal>.

MRI and Neuropsychological Data

For each subject of **Table 1**, and for each time point (*24 months before stable diagnosis*, *18 months before stable diagnosis*, *12 months before stable diagnosis*, and *time-zero point of stable diagnosis*), structural MR images were downloaded from the ADNI data repository. According to the ADNI acquisition protocol (Jack et al., 2008), examinations were performed at 1.5 T using a T1-weighted sequence. We considered MR images that had undergone the following preprocessing steps: (1) 3D gradwarp correction for geometry correction caused by gradient non-linearity (Jovicich et al., 2006), and (2) B1 non-uniformity correction for intensity correction caused by non-uniformity (Narayana et al., 1988). These preprocessing steps help improving the standardization among MR images from different MR

sites and different platforms. MR images were downloaded in 3D NIfTI format. A further processing procedure was then performed on the downloaded images, this procedure consisting in: (1) image re-orientation; (2) cropping; (3) skull-stripping; (4) image normalization to the MNI standard space by means of co-registration to the MNI template (MNI152 T1 1 mm brain) (Grabner et al., 2006; O'Hanlon et al., 2013). MR images were then segmented into Gray Matter (GM) and White Matter (WM) tissue probability maps, and smoothed using an isotropic Gaussian kernel with Full Width at Half Maximum (FWHM) ranging from 2 to 12 mm³, with a step of 2 mm³. After this phase, all MR images (whole-brain, GM and WM) resulted to be of size 121 × 145 × 121 voxels. The whole process was performed using the VMB8 software package installed on the Matlab platform (Matlab R2016b, The MathWorks). MRI volumes were visually inspected for checking homogeneity and absence of artifacts both before and after the pre-processing step.

Neuropsychological data were also obtained for each subject and for each time point from the ADNI data repository. Neuropsychological data included both scores and subscores of seven neuropsychological tests, namely the Functional Assessment Questionnaire (FAQ), the Clock Test, the Rey Auditory Verbal Learning Test (AVLT), the Digit Span (DS), the Category Fluency Tests (Animals and Vegetables), the Trail Making Test A-B (TMT A-B), and the Boston Naming Test (BNT). The full list of neuropsychological scores and subscores used in this study is reported in the Supplementary Table S1. All scores and subscores underwent a z-score normalization before being fed into the classification algorithm.

The Classification

For each subject of **Table 1**, and for each time point, T1-weighted structural MR images and neuropsychological scores (and subscores) were used as input data of an automatic binary classifier to discriminate the two groups of subjects: (CN + sMCI) vs. (pMCI + AD).

For this purpose we used an AI system based on a supervised ML algorithm, tailored to learn from MRI images the prediction model to classify different diagnostic AD groups (Salvatore et al., 2015b).

The whole procedure is detailed in the following Sub-sections and consists into: extraction of features from the three different segmented MR images (whole-brain, GM or WM); ranking of features extracted from MR images; ranking of normalized neuropsychological scores and sub-scores; classification of subjects using the extracted and ranked features, further selected according to their ranking through a wrapper procedure. This procedure is repeated for different combinations of selected

TABLE 1 | Demographic and clinical characteristics of the subjects considered in this study.

| Group type (stable diagnosis) | # Subjects | Age mean ± std. [range] | Gender #M/#F (%) |
|----------------------------------|------------|----------------------------|---------------------|
| CN or sMCI | 100 | 74.8 ± 6.4 [58.0–87.7] | 55% |
| pMCI or AD | 100 | 74.7 ± 7.1 [55.3–88.4] | 54% |

features, and the classifier is optimized on that combination showing the best classification performance (wrapper feature selection and optimization of classification).

Feature Extraction and Ranking

Feature extraction and feature ranking were performed to reduce the number of features to be handled by the classification algorithm, to remove the noisy features while keeping the ones relevant for group discrimination, and to reduce redundancy in the dataset. Thus, this step allowed an enhancement of the performance of the ML classifier while reducing computational costs.

A Principal Component Analysis (PCA) was implemented to perform feature extraction from the MRI volumes (López et al., 2011; Salvatore et al., 2015a). In particular, this method consists in applying an orthogonal transformation to the original set of variables in order to obtain a new (smaller) set of orthogonal variables called principal components. These new variables define a subspace, called the PCA subspace. The original dataset is then projected onto the PCA subspace, this operation resulting in a smaller set of features which are referred to as PCA coefficients and which can be used to replace the original dataset. This new dataset of PCA coefficients maximizes the variance of the dataset, under the constraint of orthogonality among the extracted variables. The number of extracted features cannot be higher than the value of the smaller dimension of the original dataset – 1. In our case, being the dimension of the dataset equal to $S \times N$, where S is the number of samples (200) and N the number of features (MRI voxels + neuropsychological features, $> 10^6$), then the number of extracted PCA coefficients will be at most 199.

Feature ranking was applied to PCA coefficients extracted from MR images, as well as to neuropsychological scores and sub-scores. FDR was implemented to perform feature ranking, which aims at sorting features according to their class-discriminatory power. This index was computed for each variable as follows:

$$FDR = \frac{(\mu_A - \mu_B)^2}{\sigma_A^2 + \sigma_B^2} \quad (1)$$

where the numerator expresses the squared difference between the mean of that variable in class A and class B, while the denominator expresses the sum of the squared variances of that variable in class A and in class B.

A second independent feature-extraction technique based on Partial Least Squares (PLS) (Wold et al., 1984; Ramírez et al., 2010; Khedher et al., 2015) was implemented. The approach used in PLS is similar to the one used in PCA. However, differently from PCA, this technique involves the concurrent use of information from both the set X of observed variables (the original dataset itself) and the corresponding set T of diagnostic labels. Specifically, PLS consists in computing orthogonal vectors (also in this case called components) by maximizing the covariance between the two sets of variables X and T . The original variables are then projected onto the new space spanned by the computed orthogonal vectors. These

projections are then used as input features for the classification system.

The feature-extraction-and-ranking technique based on PCA+FDR and the feature-extraction technique based on PLS were implemented independently from each other. The performances of the classifier implemented using these two techniques were then compared.

The Classifier

A Support Vector Machine (SVM) was used as a binary classifier (Cortes and Vapnik, 1995). The SVM algorithm was able to construct a predictive model based on a set of features from subjects with known stable diagnosis, called training dataset. This predictive model was then used to automatically classify new subjects (with unknown diagnosis) as belonging to one of the two diagnostic classes.

The predictive model computed by SVM was the one that maximized the margin between the two diagnostic classes, represented by a hyper-plane whose analytical form is given by:

$$y(x) = \sum_{n=1}^N w_n \bullet t_n \bullet k(x, x_n) + b \quad (2)$$

Here N is the number of subjects in the training set, w_n is the weight assigned by SVM to each subject n in the training set during the training phase, t_n represents the diagnosis of the subject n of the training set, $k(x, x_n)$ is the kernel function, and b is a threshold parameter.

In our analyses, we implemented a linear kernel SVM on the Matlab platform (R2016b, The MathWorks), also including algorithms from the biolearning toolbox of Matlab.

Wrapper Feature Selection, Optimization of Classification, Performance Evaluation

In order to find the best configuration of parameters for the classification, a wrapper feature selection and optimization of classification was performed. Specifically, the features to be selected were the MRI features extracted and ranked using PCA and FDR, and the neuropsychological scores and sub-scores normalized and ranked using FDR. The parameters to be optimized were only related to the MR image preprocessing, and they included the tissue probability map (whole-brain, GM or WM), and the FWHM of the smoothing kernel (FWHM = 2, 4, 6, 8, 10, and 12 mm³ or no smoothing).

Wrapper feature selection and optimization were performed using a fivefold Nested-Cross-Validation (Nested CV) approach (Varma and Simon, 2006). In this approach, the original dataset (100 subjects with CN or sMCI and 100 subjects with AD or pMCI) was split into 5 subsets of equal size: 4/5 subsets were used in an inner training-and-validation loop to perform feature selection and parameter optimization; the remaining 1/5 subset was then used in an outer test loop for the performance evaluation of the classifier. This procedure was repeated five times, until all subsets were used once for testing in the outer loop.

For each round, the set of selected features and optimal parameters was estimated in the inner loop as the one that

maximized the accuracy of classification. For each round, the performance was estimated in the outer loop in terms of accuracy, sensitivity, and specificity of classification. Mean accuracy, sensitivity and specificity was calculated averaging across all 5 rounds.

Given that the number of subjects in the whole dataset was 200 (i.e., 100 CN + sMCI and 100 pMCI + AD), for each round of nested CV the number of subjects used to train the classifier was 128, the number of subjects used to optimize the classifier was 32 (inner loop), and the number of subjects used to evaluate the performance of the classifier was 40 (outer loop).

The whole process was performed for each time point (24 months before stable diagnosis, 18 months before stable diagnosis, and 12 months before stable diagnosis).

In order to assess the statistical significance of each performance metric (accuracy, sensitivity, and specificity of classification), we performed a permutation test. Specifically, the classifier was run as described above, but the labels were computed as a random permutation of the original label set. This procedure was repeated for a total of 1000 iterations. A p -value indicating the statistical significance of each performance metric was then calculated as the fraction of the total number of iterations for which the performance (accuracy, sensitivity, or specificity, respectively) resulted to be greater than or equal to the performance observed using the original labels.

MRI and Neuropsychological Predictors

A three-dimensional map of voxel-based intensity distribution of MRI differences between (CN + sMCI) and (pMCI + AD) was generated for each round of the inner training-and-validation loop. The map was created for the set of selected features and optimal parameters obtained using the PCA+FDR feature-extraction-and-ranking technique. The maps generated during the 5 rounds of nested CV were then averaged in a single final map.

The importance of each voxel was computed as in our previous papers (Cerasa et al., 2015; Salvatore et al., 2015b) based on the predictive model generated by SVM. Specifically, during the training phase, SVM assigns a weight to each sample in the training set corresponding to the importance of that sample in defining the predictive model. By multiplying each sample of the training set by the corresponding weight, and by adding resulting weighted samples on a voxel-basis, it is possible to generate a three-dimensional map of the weights of each voxel. Furthermore, the method proposed by Haufe et al. (2014) to compute activation patterns in backward models was applied in order to ensure the correct interpretation of the weights.

Voxel-based maps were then normalized in intensity (to a range between 0 and 1) and superimposed on a standard stereotactic brain using a proper color scale. This procedure was performed for each time point (24 months before stable diagnosis, 18 months before stable diagnosis, and 12 months before stable diagnosis) (Cerasa et al., 2015; Salvatore et al., 2015b).

The most frequent neuropsychological scores and subscores among those selected in all rounds were also identified. Also in this case, these results were obtained for the classifier

implemented using the PCA+FDR feature-extraction-and-ranking technique. These features were sorted in descending order according to their frequency. The features occurring with a higher frequency than 5% were shown as best predictors.

RESULTS

The Classification

Classification results when using PCA+FDR as feature-extraction-and-ranking technique are shown in **Table 2** for the classification of (CN + sMCI) vs. (pMCI + AD). Using only MRI data, accuracy, sensitivity, and specificity of the classification were 0.72 ± 0.08 , 0.69 ± 0.12 , and 0.75 ± 0.08 , respectively, at the time point 24 months before stable diagnosis; 0.77 ± 0.05 , 0.78 ± 0.07 , and 0.76 ± 0.10 at the time point 18 months before stable diagnosis; 0.75 ± 0.08 , 0.79 ± 0.14 , and 0.71 ± 0.11 at the time point 12 months before stable diagnosis. As a benchmark, we also measured the performance of the classifier in discriminating (CN + sMCI) vs. (pMCI + AD) at the time-zero point of stable diagnosis (that is, when all pMCI had manifested their progression to AD). In this case, accuracy, sensitivity and specificity resulted to be 0.79 ± 0.08 , 0.83 ± 0.14 , and 0.75 ± 0.10 , respectively. The performances of the proposed method result to be statistically significant as assessed by means of permutation tests ($p < 0.001$). On the other side, no statistical difference was found among the performance obtained at the four different time points ($p = 0.51$, one-way ANOVA). The p -values (multiple comparisons for one-way ANOVA) for all the possible binary combinations of time points are reported in the Supplementary Table S2.

When using MRI and neuropsychological data in combination, accuracy, sensitivity, and specificity were 0.85 ± 0.05 , 0.83 ± 0.09 , and 0.87 ± 0.06 , respectively, at the time point 24 months before stable diagnosis; 0.85 ± 0.09 , 0.86 ± 0.11 , and 0.83 ± 0.17 at the time point 18 months before stable diagnosis; 0.87 ± 0.06 , 0.86 ± 0.11 , and 0.87 ± 0.03 at the time point 12 months before stable diagnosis. Accuracy, sensitivity and specificity at the time-zero point of stable diagnosis were 0.92 ± 0.01 , 0.91 ± 0.04 , and 0.93 ± 0.03 , respectively. The performances of the proposed method result to be statistically significant as assessed by means of permutation tests ($p < 0.001$). On the other side, no statistical difference was found among the performance obtained at the four different time points ($p = 0.20$, one-way ANOVA). The p -values (multiple comparisons for one-way ANOVA) for all the possible binary combinations of time points are reported in the Supplementary Table S3.

Furthermore, when comparing -at different time points- the accuracy of classification obtained using MRI and neuropsychological data in combination with respect to the one obtained using MRI alone, the combined approach resulted to perform statistically better -at the 5% significance level- than the single-modality approach at the time points of 24 months before stable diagnosis ($p = 0.01$), 12 months before stable diagnosis ($p = 0.03$), and at the stable-diagnosis time point ($p = 0.01$). No statistical difference was found at the time point of 18 months before stable diagnosis ($p = 0.15$).

TABLE 2 | Classification performance in terms of accuracy, sensitivity, and specificity for (CN + sMCI) vs. (pMCI + AD) at the considered time points, using MR images alone or coupled with neuropsychological measures, with PCA+FDR as feature-extraction-and-ranking technique.

| | 24 m before stable diagnosis | 18 m before stable diagnosis | 12 m before stable diagnosis | Stable-diagnosis time point |
|-------------------------------|------------------------------|------------------------------|------------------------------|-----------------------------|
| MRI | | | | |
| Accuracy | 0.72 ± 0.08 | 0.77 ± 0.05 | 0.75 ± 0.08 | 0.79 ± 0.08 |
| Sensitivity | 0.69 ± 0.12 | 0.78 ± 0.07 | 0.79 ± 0.14 | 0.83 ± 0.14 |
| Specificity | 0.75 ± 0.08 | 0.76 ± 0.10 | 0.71 ± 0.11 | 0.75 ± 0.10 |
| MRI + Neuropsychological data | | | | |
| Accuracy | 0.85 ± 0.05 | 0.85 ± 0.09 | 0.87 ± 0.06 | 0.92 ± 0.01 |
| Sensitivity | 0.83 ± 0.09 | 0.86 ± 0.11 | 0.86 ± 0.11 | 0.91 ± 0.04 |
| Specificity | 0.87 ± 0.06 | 0.83 ± 0.17 | 0.87 ± 0.03 | 0.93 ± 0.03 |

The performance of the classifier at the time point of the stable diagnosis is also shown.

TABLE 3 | Classification performance in terms of accuracy, sensitivity, and specificity for (CN + sMCI) vs. (pMCI + AD) at the considered time points, using MR images alone or coupled with neuropsychological measures, with PLS as feature-extraction technique.

| | 24 m before stable diagnosis | 18 m before stable diagnosis | 12 m before stable diagnosis | Stable-diagnosis time point |
|-------------------------------|------------------------------|------------------------------|------------------------------|-----------------------------|
| MRI | | | | |
| Accuracy | 0.79 ± 0.07 | 0.81 ± 0.04 | 0.81 ± 0.05 | 0.82 ± 0.04 |
| Sensitivity | 0.79 ± 0.07 | 0.81 ± 0.07 | 0.83 ± 0.08 | 0.82 ± 0.07 |
| Specificity | 0.78 ± 0.08 | 0.81 ± 0.07 | 0.79 ± 0.05 | 0.81 ± 0.04 |
| MRI + Neuropsychological data | | | | |
| Accuracy | 0.81 ± 0.07 | 0.83 ± 0.12 | 0.84 ± 0.06 | 0.85 ± 0.05 |
| Sensitivity | 0.82 ± 0.08 | 0.83 ± 0.10 | 0.86 ± 0.07 | 0.87 ± 0.09 |
| Specificity | 0.80 ± 0.11 | 0.83 ± 0.18 | 0.82 ± 0.10 | 0.83 ± 0.04 |

The performance of the classifier at the time point of the stable diagnosis is also shown.

Classification results obtained when using PLS as feature extraction technique are shown in **Table 3**. Using only MRI data, accuracy, sensitivity and specificity of the classification were 0.79 ± 0.07 , 0.79 ± 0.07 , and 0.78 ± 0.08 , respectively, at the time point 24 months before stable diagnosis; 0.81 ± 0.04 , 0.81 ± 0.07 , and 0.81 ± 0.07 at the time point 18 months before stable diagnosis; 0.81 ± 0.05 , 0.83 ± 0.08 , and 0.79 ± 0.05 at the time point 12 months before stable diagnosis. The benchmark performance of the classifier at the time-zero point of stable diagnosis was 0.82 ± 0.04 accuracy, 0.82 ± 0.07 sensitivity and 0.81 ± 0.04 specificity. The performances of the proposed method resulted to be statistically significant as assessed by means of permutation tests ($p < 0.001$). No statistical difference was found among the performance obtained at the four different time points ($p = 0.76$ for accuracy, one-way ANOVA). The p -values (multiple comparisons for one-way ANOVA) for all the possible binary combinations of time points are reported in the Supplementary Table S4.

When using a combination of MRI and neuropsychological data, accuracy, sensitivity and specificity were 0.81 ± 0.07 , 0.82 ± 0.08 , and 0.80 ± 0.11 , respectively, at the time point 24 months before stable diagnosis; 0.83 ± 0.12 , 0.83 ± 0.10 , and 0.83 ± 0.18 at the time point 18 months before stable diagnosis; 0.84 ± 0.06 , 0.86 ± 0.07 , and 0.82 ± 0.10 at the time point 12 months before stable diagnosis. The benchmark performance of the classifier in terms of accuracy, sensitivity and specificity at the time-zero point of stable diagnosis was 0.85 ± 0.05 , 0.87 ± 0.09 , and 0.83 ± 0.04 , respectively. The performances of the proposed

method result to be statistically significant as assessed by means of permutation tests ($p < 0.001$). No statistical difference was found among the performance obtained at the four different time points ($p = 0.88$ for accuracy, one-way ANOVA). The p -values (multiple comparisons for one-way ANOVA) for all the possible binary combinations of time points are reported in the Supplementary Table S5.

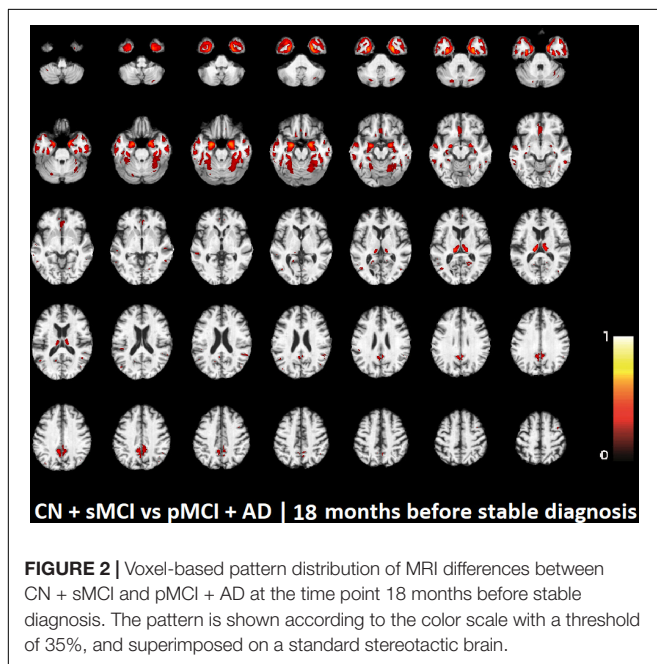
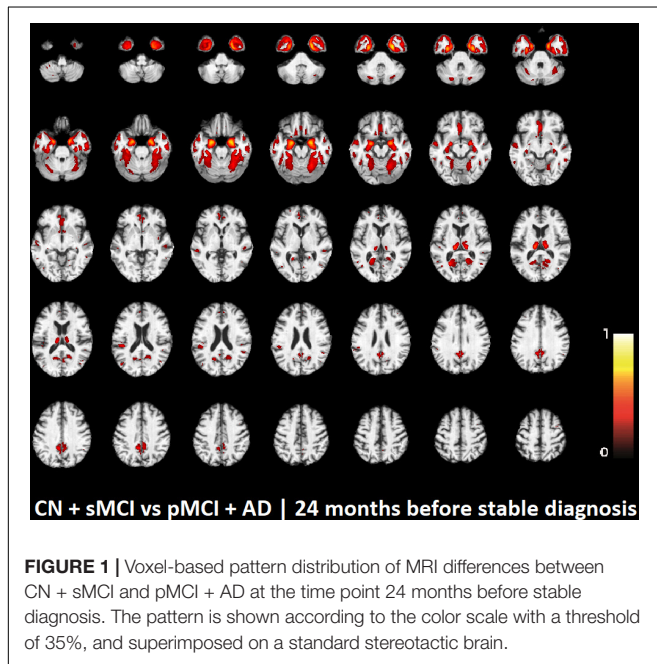
Furthermore, when comparing –at different time points– the accuracy of classification obtained using MRI and neuropsychological data in combination with respect to the one obtained using MRI alone, no statistical difference was observed ($p = 0.23$ at the time point of 24 months before stable diagnosis; $p = 0.65$ at the time point of 18 months before stable diagnosis; $p = 0.11$ at the time point of 12 months before stable diagnosis; $p = 0.08$ at the stable-diagnosis time point).

Making a pairwise comparison (paired-sample t -test) between the performance obtained using PCA+FDR vs. PLS (for each time point and for each domain), results show that –at the 5% significance level– the classifier implemented using PLS performed statistically better (in terms of accuracy) than the one implemented using PCA+FDR at the time points of 24 and 18 months before stable diagnosis when using MRI alone ($p = 0.03$ in both cases). A comprehensive table showing all pairwise p -values can be found in Supplementary Table S6.

MRI and Neuropsychological Predictors

The voxel-based pattern distribution of MRI differences found as results of classification between CN + sMCI

and $pMCI + AD$ are shown in **Figures 1–3**, for the three considered time points, respectively (i.e., *24 months before stable diagnosis*, *18 months before stable diagnosis*, and *12 months before stable diagnosis*). The voxel-based pattern distribution of MRI differences at the time-zero point of stable diagnosis is also shown in **Figure 4**. All patterns were shown according to the color scale with a threshold of 35%, and superimposed on a standard stereotactic brain in order to allow a better localization of the brain regions identified by the classifier.



Similarly, the best neuropsychological predictors and corresponding status/domain/subdomain found for the classification of (CN + sMCI) vs. (pMCI + AD) for the considered time-points are reported in **Table 4**. Findings are sorted in descending order according to their frequency. The complete list of best neuropsychological predictors with the corresponding names as reported in the ADNI data repository can be found in Supplementary Table S7.

DISCUSSION

The main finding of our work was that, using structural T1-weighted MRI brain studies and specific neuropsychological measures, our classifier was able to identify mild-AD patients who need treatments 24 months before AD definite diagnosis with an 85% accuracy, 83% sensitivity, and 87% specificity (see **Table 2**, when considering the method implemented using PCA+FDR). More interestingly, the performance obtained by our multi-modal classifier in distinguishing normal subjects (or stable MCI) from patients who will evolve to AD 24 months before stable diagnosis is comparable ($p > 0.2$) to the ones obtained at 18, 12 months before stable diagnosis and, even more important, to the one obtained at the time of definite diagnosis. Furthermore, the combined classification approach model outperformed the other classification considered in this study using single MRI data (72% classification accuracy, 69% sensitivity, and 75% specificity) (**Table 2**, $p < 0.05$, when considering the method implemented using PCA+FDR).

Although the discrimination of (CN + sMCI) vs. (pMCI + AD) is not common in the literature, our results can be compared with the classification performance of studies focused on predicting the conversion to Alzheimer's dementia. These studies usually limit their attention to the binary classification of $pMCI$ vs. $sMCI$. In a recent review considering 30 studies applying ML for the diagnosis of AD using only structural MRI (Salvatore et al., 2015a), the mean classification accuracy in discriminating $pMCI$ vs. $sMCI$ was found to be 0.66 ± 0.11 . Another study tried to distinguish AD patients from stable MCI patients using only structural MRI features (Diciotti et al., 2012). A classification accuracy of 0.74 was reported (0.72 sensitivity, 0.77 specificity), although they used a private cohort of 21 mild AD and 30 MCI patients, and the gold-standard diagnosis was not based on follow-up examinations. Some other studies tried to automatically classify $pMCI$ vs. $sMCI$ using only MRI features (e.g., Cui et al., 2011; Koikkalainen et al., 2012; Ye et al., 2012; Casanova et al., 2013; Peters et al., 2014; Runtti et al., 2014; Dukart et al., 2015; Eskildsen et al., 2015; Moradi et al., 2015; Ritter et al., 2015; Salvatore et al., 2015b; Nanni et al., 2016), with a classification accuracy ranging from 0.51 to 0.75.

To the best of our knowledge, this is one of the few works able to answer the question whether a multidisciplinary classification model coupling cognitive, functional and behavioral measures with structural MRI brain studies is better than a model based only on structural MRI. Four studies attempted the task of classifying $pMCI$ vs. $sMCI$ using both structural-MRI features alone and in combination with neuropsychological measures

(Cui et al., 2011; Runtti et al., 2014; Dukart et al., 2015; Moradi et al., 2015). The classification accuracy of these studies ranges from 0.62 to 0.75 when using structural MRIs alone, and from 0.62 to 0.82 when using both structural MRIs and neuropsychological measures, showing a slight improvement (the mean intra-study improvement was 0.06 ± 0.04).

Another challenging finding of our study was that patterns of morphological abnormalities localized in the temporal pole and medial-temporal cortex might be considered as biomarkers of clinical progression and evolution (Figures 1–4). These regions

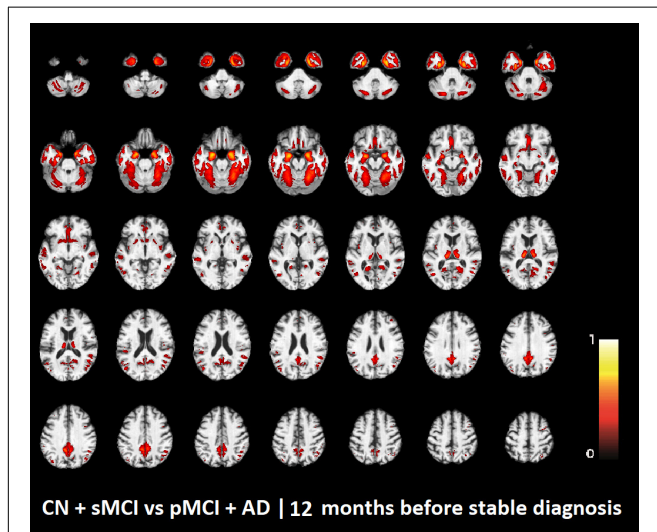


FIGURE 3 | Voxel-based pattern distribution of MRI differences between CN + sMCI and pMCI + AD at the time point 12 months before stable diagnosis. The pattern is shown according to the color scale with a threshold of 35%, and superimposed on a standard stereotactic brain.

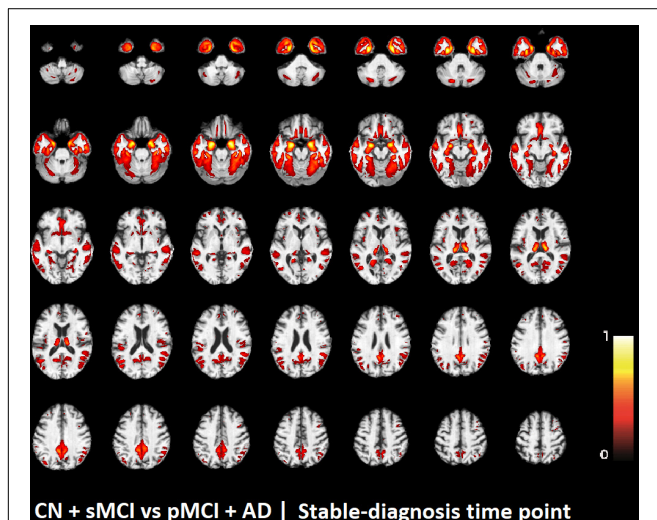


FIGURE 4 | Voxel-based pattern distribution of MRI differences between CN + sMCI and pMCI + AD at the time-zero point of stable diagnosis. The pattern is shown according to the color scale with a threshold of 35%, and superimposed on a standard stereotactic brain.

can be already observed at the time point of *24 months before stable diagnosis* (Figure 1). When considering the subsequent time points (Figures 2–4), the voxel-based pattern distribution of MRI-related neurodegeneration is similar to that one at 24 months before stable diagnosis, but progressively more extended, which could be a consequence of a more advanced process of structural neurodegeneration. There is an increasing interest proven by literature in understanding progression-related brain changes using structural MRI, describing an association between progression and atrophy, especially of the parietal and posterior cingulate regions, extending into the precuneus and medial temporal regions including hippocampus, amygdala, and entorhinal cortex. This pattern of progression-atrophy association is even evident at mild stages of cognitive impairment. The purpose of our work is out from explaining mechanisms behind the structural pattern distribution related to MRI images of different stages of disease progression. However, the progressive pattern seems to be consistent with Braak pathological studies (Braak and Braak, 1991), showing that during the development of AD pathology, tau tangles increase, associated with synapse loss and neurodegeneration.

Finally, we demonstrated that some cognitive, functional, and behavioral measures emerged as best predictors for AD progression. These include measures of functional abilities, memory and learning, working memory, language, visuoconstructional reasoning, and complex attention (see Table 4). More specifically, the best neuropsychological predictors for the classification of (CN + sMCI) vs. (pMCI + AD) at the time point of *24 months before stable diagnosis* include measures of functional abilities, memory and learning, working memory, and language. When considering the subsequent time points, involved domains are similar to the ones at 24 months before stable diagnosis. Interestingly, some of the sub-scores obtained through the administration of the FAQ (domain: functional abilities) and AVLT (domain: memory and learning) are always selected as best neuropsychological predictors at all the considered time points. Moreover, it must be noted that the best neuropsychological predictors at the time point of *stable diagnosis* include only measures from these two tests, which could be a consequence of a more advanced impairment in these two domains. Neuropsychological assessment can be time intensive, and the experience of practitioners can impact on the reliability and efficiency of the assessment. Our results can help the work of clinicians in optimizing the choice of cognitive tests to be administered at no costs for effectiveness. In a previous study of our group, Battista et al. (2017) demonstrated that it is possible to use a selected subset of neuropsychological measures to automatically diagnose AD patients with an accuracy of 90%.

It should be underlined that -in the present study- most of the best neuropsychological predictors at the time point of *24 months before stable diagnosis* are components of the AVLT or partial scores of FAQ related to learning and verbal episodic memory or prospective memory. These findings may confirm that the best neuropsychological predictors of conversion from amnesic MCI to AD are tests of episodic memory, as recently pointed out by Gainotti et al. (2014). Furthermore, also in the above-cited paper by Battista et al. (2017) the subset of selected neuropsychological

TABLE 4 | Best Neuropsychological predictors and corresponding status/domain/subdomain found for the classification of (CN + sMCI) vs. (pMCI + AD).

| Time point | Neuropsychological predictor | Status/domain/subdomain of predictor |
|-------------------------------------|---|--------------------------------------|
| 24 months before stable diagnosis | Ability in remembering appointments, family occasions, holidays, medications in FAQ | Functional abilities |
| | Ability in writing checks, paying bills, or balancing checkbook in FAQ | Functional abilities |
| | Ability in assembling tax records, business affairs in FAQ | Functional abilities |
| | Total score of trial 5 in AVLT | Memory and learning |
| | Ability in keeping track of current events in FAQ | Functional abilities |
| | Total intrusions of trial 1 in AVLT | Memory and learning |
| | Correct answers in the Backwards task in Digit-Span Test | Working memory |
| | Correct answers in Vegetables task in Category Fluency Test | Language |
| | Correct answers after a 30-min delay in AVLT | Memory and learning |
| | | |
| 18 months before stable diagnosis | Ability in writing checks, paying bills, or balancing checkbook in FAQ | Functional abilities |
| | Ability in remembering appointments, family occasions, holidays, medications in FAQ | Functional abilities |
| | Total score of trial 3 in AVLT | Memory and learning |
| | Total score of trial 5 in AVLT | Memory and learning |
| | Total score of trial 6 in AVLT | Memory and learning |
| | Ability in assembling tax records, business affairs in FAQ | Functional abilities |
| | Ability in traveling, driving, or arranging to take public transportation in FAQ | Functional abilities |
| | Presence of the two hands in CLOCK test | Visuoconstructional reasoning |
| | Ability in shopping alone for necessities in FAQ | Functional abilities |
| | Ability in keeping track of current events in FAQ | Functional abilities |
| | Total score of FAQ | Functional abilities |
| | Total of trial 4 in AVLT | Memory and learning |
| | Spontaneously given correct responses in BNT | Language |
| | Corrected responses following phonemic cues in BNT | Language |
| | Symmetry of number placement in CLOCK test | Visuoconstructional reasoning |
| | Presence of the two hands, set to ten after eleven in CLOCK test | Visuoconstructional reasoning |
| | Time to complete Part A of the test in TMT | Complex attention |
| | Time to complete Part B of the test in TMT | Complex attention |
| | Correct answers after a 30-min delay in AVLT | Memory and learning |
| | Recognition errors in AVLT | Memory and learning |
| 12 months before stable diagnosis | Ability in writing checks, paying bills, or balancing checkbook in FAQ | Functional abilities |
| | Ability in remembering appointments, family occasions, holidays, medications in FAQ | Functional abilities |
| | Total of trial 3 in AVLT | Memory and learning |
| | Number of correct responses following a phonemic cue in BNT | Language |
| | Ability in assembling tax records, business affairs in FAQ | Functional abilities |
| | Ability in shopping alone for necessities in FAQ | Functional abilities |
| | Ability in traveling, driving, or arranging to take public transportation in FAQ | Functional abilities |
| | Total score of FAQ | Functional abilities |
| | Total of trial 4 in AVLT | Memory and learning |
| | Total of trial 5 in AVLT | Memory and learning |
| | Total correct answers after a 30-min delay in AVLT | Memory and learning |
| | Total of trial 6 in AVLT | Memory and learning |
| | Ability in keeping track of current events in FAQ | Functional abilities |
| | Ability in paying attention to and understanding a TV program, book, or magazine in FAQ | Functional abilities |
| | Total score of the CLOCK test | Visuoconstructional reasoning |
| Time-zero point of stable diagnosis | Ability in writing checks, paying bills, or balancing checkbook in FAQ | Functional abilities |
| | Total score of FAQ. | Functional abilities |
| | Total of trial 4 in AVLT | Memory and learning |
| | Ability in remembering appointments, family occasions, holidays, medications in FAQ | Functional abilities |
| | Ability in paying attention to and understanding a TV program, book, or magazine in FAQ | Functional abilities |
| | Ability in traveling out of the neighborhood, driving, arranging to take public transportation in FAQ | Functional abilities |
| | Ability in assembling tax records, business affairs, or other papers in FAQ | Functional abilities |
| | Ability of the subject in preparing a balanced meal in FAQ | Functional abilities |
| | Total of trial 6 in AVLT | Memory and learning |
| | Ability in playing a game of skill such as bridge or chess, working on a hobby in FAQ | Functional abilities |
| | Correct answers after a 30-min delay in AVLT | Memory and learning |

Results are reported for the three considered time points (i.e., 24 months before stable diagnosis, 18 months before stable diagnosis, and 12 months before stable diagnosis) and for the time-zero point of stable diagnosis. Best neuropsychological predictors are sorted in descending order according to their frequency (the frequency of that measure in all loops). The status/domain/subdomain corresponding to the neuropsychological predictor is also reported.

measures able to automatically diagnose AD patients was mainly composed of measures related to episodic memory (namely, scores and subscores of AVLT, Logical Memory Test and

Alzheimer's Disease Assessment Scale-Cognitive Behavior) and measures addressing functional abilities in daily life (namely, total score and subscores of FAQ).

With respect to the numerous other ML methods proposed for the automatic classification of AD patients by means of brain MRI images (Cuingnet et al., 2011; Salvatore et al., 2015a), our approach has several points of strength.

Firstly, we validated our data on a large, multi-center independent cohort study, namely the ADNI public database. The use of large, public cohorts for training machine-learning classifiers allows a higher generalization ability than using private cohorts, which are often obtained from single-center studies. Moreover, the use of public databases is crucial for the comparison of the classification performance of different studies (Cuingnet et al., 2011), which is not recommended for studies using different private inhomogeneous cohorts. Mainly because of these reasons, in the last few years, the use of large, public data repositories is becoming more frequent in the field of ML applied to neuroimaging data, as reported in a recent review (Salvatore et al., 2015a). However, to date this is not a standard practice, and several studies still make use of private cohorts.

A second point of strength is that our algorithm requires a limited number of imaging studies to be trained, nearly a hundred studies per diagnostic class. This point is particularly important if considered with respect to the new classification approaches that are recently emerging as state-of-the-art techniques in the computer-vision community, namely deep-learning. These techniques have proven to be high performing in most automatic-classification tasks (Sharif Razavian et al., 2014), but their application in medicine, in particular in the neuroimaging field, is still limited. This is due to the requirement of at least a thousand of imaging studies per diagnostic class in order to reduce overfitting problems.

The third point of strength is the ability of our classification algorithm to return the best MRI and neuropsychological predictors, that is, the most important structural-brain patterns and neuropsychological scores for distinguishing the two diagnostic classes. Specifically, these predictors can be interpreted as early signs of the disease, and thus be used as surrogate biomarkers of AD. In the case of structural-MRI predictors, this may be particularly useful in monitoring the course of the neurodegeneration or the efficacy of a treatment.

Another advantage of our classification algorithm is that data used as input can be collected in a single examination session following routinely clinical protocols (T1-weighted MRI on 1.5T systems) and non-invasive and inexpensive measures obtained through the administration of standard neuropsychological tests.

Lastly, with respect to the use of structural MRI volumes, it must be noted that our classification algorithm does not require any interaction or pre-processing by the neuroradiologists on the original acquired images. This helps avoiding any issue arising from inter- and intra-operator inhomogeneities.

From a methodological point of view, we must underline two further points of strength. The first is the number of features used for training the classification algorithm, which was lower than the number of subjects in the two classes. This practice is useful as it prevents any curse-of-dimensionality issue. The second is the independence between neuropsychological measures used as features and measures used as gold standard to perform the original classification in the four diagnostic groups (AD, pMCI, sMCI, and CN).

This practice warrants the avoidance of double-dipping in the classification process (Kriegeskorte et al., 2009).

However, we should also recognize some limitations of our work:

Limited Generalization Ability and Reliability. Further investigations are needed in order to assess the generalization ability and reliability of our multimodal MRI/cognitive-based classifier, and its applicability at an individual subject level. Our results are based on subjects in the United States and Canada, thus validation studies including subjects from other regions worldwide are lacking. Moreover, our predictive results have been obtained by a cross-validation approach using these subjects, and this may not accurately generalize our findings to a general population. We have used an SVM classifier since it offers different advantages, for example, is particularly appropriate for non-linear and big data such as whole-brain MRI images, also in combination with data from other modalities (e.g., biological and neuropsychological data). However, in order to confirm our results, we should have used more classifiers among the variety of ML methods already validated for automatic classification of medical images, e.g., Artificial Neural Networks, Linear Discriminant Analysis, regression models, Bayesian approaches, Decision Trees, and Random Forests.

Limited Clinical Questions. In this work we developed a predictive model able to address CN and sMCI subjects to a different therapeutic option with respect to pMCI and AD subjects. Our approach cannot be used for screening patients for specific AB or tau target drug clinical trials.

Approximately 27% of subjects meeting clinical inclusion criteria for mild-AD were found A β -negative, thus, our multimodal classifier does not allow to avoid variance into analyses due to these patients. A β -negative mild-AD subjects are not expected to progress clinically on the expected trajectory, adding variance into analyses where a slowing of progression is being measured. Clinical trials of putative therapeutics for AD should use a baseline measure of brain A β or tau as an inclusion criterion, such as PET amyloid studies, even if a recent work demonstrated that measuring A β status from MRI scans in mild-AD subjects is possible and may be a useful screening tool in clinical trials (Tosun et al., 2016).

Limited Neuropsychological Predictors. Our work considered neuropsychological scores and sub-scores obtained from seven neuropsychological tests as candidate predictors. Whilst this offered a certain amount and details of information on different cognitive domains (a total of 64 scores were used as input data) as well as on behavioral and functional status, many other measures coming from other tests were excluded from our analysis only because not available for all the considered subjects. This limits our findings. A best accuracy in the prediction model could be achieved by using more neuropsychological measures (selected on the basis of their classification performance).

Limited Dynamic View of the Disease Progression. This study lacks of a dynamic view of the disease progression in terms of linking the imaging data between different time points. Although the different patterns of cerebral changes in AD/MCI over several time points have been compared in this paper, the proposed analysis was cross-sectional in nature at each

time point, thus not investigating cross-time-point relationships with the predictive models. This would be a fundamental step for advancing our knowledge about neuropathological staging of Alzheimer-related changes. However, it should be kept in mind that in the last 10 years a plethora of longitudinal studies have provided consistent evidence on the evolution of neurodegenerative changes in AD brain. Recent advances in molecular neuroimaging have greatly facilitated our ability to detect neurodegenerative pathology *in vivo*, particularly in the very early stages of AD. As recently reviewed by Sperling et al. (2014), the inexorable progression of neurodegeneration characterizing patients with AD begins well more than a decade prior to the stage of clinically detectable symptoms. Amyloid- β ($A\beta$) accumulation may be evident 20 years before the stage of dementia, whilst substantial neuronal loss became evident by the stage of MCI. The challenge in this new era of neuroimaging application on AD is to demonstrate the real role played by the first hallmark of AD: $A\beta$ accumulation. The general opinion is that $A\beta$ is necessary, but not sufficient in isolation, to predict imminent decline along the AD trajectory. For this reason, structural neuroimaging can be useful for increasing the accuracy of automated diagnostic methods. Overall $A\beta$ accumulation begins in the temporal cortex in very early AD phases, promoting dysmetabolism and neural losses. In the next phases, pathological changes move toward associative neocortex, mainly including orbitofrontal cortex, precuneus and prefrontal cortex, finally reaching the primary motor system along the AD trajectory. Our findings are thus in agreement with the well-known neurodegenerative staging of AD brain.

Limited Prediction Over the Course of Disease. In this study we were not able to establish if predicting progression to AD of MCI patients could be possible even at an earlier time than the 24 months prior to the definite diagnosis, since the number of subjects provided by ADNI with an entire multimodal set of measures and with a longer follow up than 24 months is not sufficient for training-and-classification purposes.

Our classifier has been trained on measures of cognitive impairment obtained through clinically administered neuropsychological-test predictors. Thus, with this configuration, it cannot be used for screening presymptomatic subjects. However, in principle, our classifiers could be trained even over a different set of cognitive/behavioral and functional data, measured during daily life of CN subjects in order to capture domains that are affected first by the disease, eventually combined with their MRI brain studies in order to detect very subtle brain changes and on biological CSF with proper established cut points.

As pointed out in a recent review by ADNI (Weiner et al., 2017), longitudinal studies aimed at the early diagnosis and prognosis of AD are able to increase the power of clinical trials, as they can help in the selection of trial participants likely to decline. In these studies, the use of ML algorithms has been proved effective to measure surrogate diagnostic biomarkers, especially in challenges involving MCI subjects, but have been poorly validated for detecting the power of measures of longitudinal changes over time as surrogate predictive biomarkers of the disease.

In our study we demonstrated that it is possible to predict the conversion of MCI to probable AD up to 24 months before the definite diagnosis. Although better suited to trials of treatments aiming to repair brain tissue rather than clear $A\beta$, our approach may improve the feasibility of clinical trials by reducing costs and increasing the power to detect disease progression.

In conclusions, to our knowledge, this is one of the few works able to answer the question whether a multidisciplinary classification model coupling cognitive, functional and behavioral measures with structural MRI brain studies is better than a model based on structural MRIs alone. Since T1-weighted MRI scans are acquired routinely in clinical trials for other purposes and neuropsychological assessment can be easily performed to complement routine clinical trials, our multimodal pMCI classifier might be useful as a screening tool that could be applied to reduce the number of non-progressive subjects not to be treated.

AUTHOR CONTRIBUTIONS

CS, AC, and IC conceived, designed, and drafted this work. CS and IC performed the artificial-intelligence analysis. All authors critically revised, and approved the final version and agreed to be accountable for this work.

FUNDING

This work was supported by the CNR Research Project “Aging: Molecular and Technological Innovations for Improving the Health of the Elderly” No. DSB.AD009.001 and Activity No. DSB.AD009.001.043. Data collection and sharing for this project was funded by the Alzheimer’s Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer’s Association; Alzheimer’s Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd. and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute

for Research and Education, and the study is coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

REFERENCES

- Battista, P., Salvatore, C., and Castiglioni, I. (2017). Optimizing neuropsychological assessments for cognitive, behavioral, and functional impairment classification: a machine learning study. *Behav. Neurol.* 2017:1850909. doi: 10.1155/2017/1850909
- Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. New York, NY: Springer, 98–108.
- Braak, H., and Braak, E. (1991). Neuropathological staging of Alzheimer-related changes. *Acta Neuropathol.* 82, 239–259. doi: 10.1007/BF00308809
- Casanova, R., Hsu, F.-C., Sink, K. M., Rapp, S. R., Williamson, J. D., Resnick, S. M., et al. (2013). Alzheimer's disease risk assessment using large-scale machine learning methods. *PLoS One* 8:e77949. doi: 10.1371/journal.pone.0077949
- Cerasa, A., Castiglioni, I., Salvatore, C., Funaro, A., Martino, I., Alfano, S., et al. (2015). Biomarkers of eating disorders using support vector machine analysis of structural neuroimaging data: preliminary results. *Behav. Neurol.* 2015:924814. doi: 10.1155/2015/924814
- Cortes, C., and Vapnik, V. (1995). Support-vector networks. *Mach. Learn.* 20, 273–297. doi: 10.1023/A:1022627411411
- Cui, Y., Liu, B., Luo, S., Zhen, X., Fan, M., Liu, T., et al. (2011). Identification of conversion from mild cognitive impairment to Alzheimer's disease using multivariate predictors. *PLoS One* 6:e21896. doi: 10.1371/journal.pone.0021896
- Cuingnet, R., Gerardin, E., Tessier, J., Auzias, G., Lehericy, S., Habert, M. O., et al. (2011). Automatic classification of patients with Alzheimer's disease from structural MRI: a comparison of ten methods using the ADNI database. *Neuroimage* 56, 766–781. doi: 10.1016/j.neuroimage.2010.06.013
- Dementia Statistics (2015). *Alzheimer's Disease International*. Available at: <http://www.alz.co.uk/research/statistics>
- Diciotti, S., Ginestroni, A., Bessi, V., Giannelli, M., Tessa, C., Bracco, L., et al. (2012). Identification of mild Alzheimer's disease through automated classification of structural MRI features. *Conf. Proc. IEEE Eng. Med. Biol. Soc.* 2012, 428–431. doi: 10.1109/EMBC.2012.6345959
- Dubois, B., Feldman, H. H., Jacova, C., DeKosky, S. T., Barberger-Gateau, P., Cummings, J., et al. (2007). Research criteria for the diagnosis of Alzheimer's disease: revising the NINCDS-ADRDA criteria. *Lancet Neurol.* 6, 734–746. doi: 10.1016/S1474-4422(07)70178-3
- Dukart, J., Sambataro, F., and Bertolino, A. (2015). Accurate prediction of conversion to Alzheimer's disease using imaging, genetic, and neuropsychological biomarkers. *J. Alzheimers Dis.* 49, 1143–1159. doi: 10.3233/JAD-150570
- Eskildsen, S. F., Coupe, P., Fonov, V. S., Pruessner, J. C., and Collins, D. L. (2015). Structural imaging biomarkers of Alzheimer's disease: predicting disease progression. *Neurobiol. Aging* 36(Suppl. 1), S23–S31. doi: 10.1016/j.neurobiolaging.2014.04.034
- Folstein, M. F., Folstein, S. E., and McHugh, P. R. (1975). "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *J. Psychiatr. Res.* 12, 189–198. doi: 10.1016/0022-3956(75)90026-6
- Gainotti, G., Quaranta, D., Vita, M. G., and Marra, C. (2014). Neuropsychological predictors of conversion from mild cognitive impairment to Alzheimer's disease. *J. Alzheimers Dis.* 38, 481–495.
- Grabner, G., Janke, A. L., Budge, M. M., Smith, D., Pruessner, J., and Collins, D. L. (2006). Symmetric atlas and model based segmentation: an application to the hippocampus in older adults. *Med. Image Comput. Comput. Assist. Interv.* 9, 58–66.
- Haufe, S., Meinecke, F., Görgen, K., Dähne, S., Haynes, J. D., Blankertz, B., et al. (2014). On the interpretation of weight vectors of linear models in multivariate neuroimaging. *Neuroimage* 87, 96–110. doi: 10.1016/j.neuroimage.2013.10.067
- Huys, Q. J., Maia, T. V., and Frank, M. J. (2016). Computational psychiatry as a bridge from neuroscience to clinical applications. *Nat. Neurosci.* 19, 404–413. doi: 10.1038/nn.4238
- Jack, C. R., Bernstein, M. A., Fox, N. C., Thompson, P., Alexander, G., Harvey, D., et al. (2008). The Alzheimer's disease neuroimaging initiative (ADNI): MRI methods. *J. Magn. Reson. Imaging* 27, 685–691. doi: 10.1002/jmri.21049
- Jovicich, J., Czanner, S., Greve, D., Haley, E., Van Der Kouwe, A., Gollub, R., et al. (2006). Reliability in multi-site structural MRI studies: Effects of gradient non-linearity correction on phantom and human data. *Neuroimage* 30, 436–443. doi: 10.1016/j.neuroimage.2005.09.046
- Khan, A., Corbett, A., and Ballard, C. (2017). Emerging treatments for Alzheimer's disease for non-amyloid and non-tau targets. *Expert Rev. Neurother.* 17, 683–695. doi: 10.1080/14737175.2017.1326818
- Khedher, L., Ramirez, J., Górriz, J. M., Brahim, A., Segovia, F., and The Alzheimer's Disease Neuroimaging Initiative (2015). Early diagnosis of Alzheimer's disease based on partial least squares, principal component analysis and support vector machine using segmented MRI images. *Neurocomputing* 151, 139–150. doi: 10.1016/j.neucom.2014.09.072
- Koikkalainen, J., Pölönen, H., Mattila, J., van Gils, M., Soininen, H., Lötjönen, J., et al. (2012). Improved classification of Alzheimer's disease data via removal of nuisance variability. *PLoS One* 7:e31112. doi: 10.1371/journal.pone.0031112
- Kriegeskorte, N., Simmons, W. K., Bellgowan, P. S., and Baker, C. I. (2009). Circular analysis in systems neuroscience – the dangers of double dipping. *Nat. Neurosci.* 12, 535–540. doi: 10.1038/nn.2303
- López, M., Ramírez, J., Górriz, J. M., Álvarez, I., Salas-Gonzalez, D., Segovia, F., et al. (2011). Principal component analysis-based techniques and supervised classification schemes for the early detection of Alzheimer's disease. *Neurocomputing* 74, 1260–1271. doi: 10.1016/j.neucom.2010.06.025
- McKhann, G., Drachman, D., Folstein, M., Katzman, R., Price, D., and Stadlan, E. M. (1984). Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group* under the auspices of department of health and human services task force on Alzheimer's disease. *Neurology* 34, 939–939. doi: 10.1212/WNL.34.7.939
- Mitchell, A. J., and Shiri-Feshki, M. (2009). Rate of progression of mild cognitive impairment to dementia—meta-analysis of 41 robust inception cohort studies. *Acta Psychiatr. Scand.* 119, 252–265. doi: 10.1111/j.1600-0447.2008.01326.x
- Moradi, E., Pepe, A., Gaser, C., Huttunen, H., and Tohka, J. (2015). Machine learning framework for early MRI-based Alzheimer's conversion prediction in MCI subjects. *Neuroimage* 104, 398–412. doi: 10.1016/j.neuroimage.2014.10.002
- Morris, J. C. (1993). The clinical dementia rating (CDR): current version and scoring rules. *Neurology* 43, 2412–2412. doi: 10.1212/WNL.43.11.2412-a
- Nanni, L., Salvatore, C., Cerasa, A., and Castiglioni, I. (2016). Combining multiple approaches for the early diagnosis of Alzheimer's disease. *Pattern Recognit. Lett.* 84, 259–266. doi: 10.1016/j.patrec.2016.10.010
- Narayana, P. A., Brey, W. W., Kulkarni, M. V., and Sievenpiper, C. L. (1988). Compensation for surface coil sensitivity variation in magnetic resonance imaging. *Magn. Reson. Imaging* 6, 271–274. doi: 10.1016/0730-725X(88)90401-8
- O'Hanlon, E., Newell, F. N., and Mitchell, K. J. (2013). Combined structural and functional imaging reveals cortical deactivations in grapheme-color synaesthesia. *Front. Psychol.* 4:755. doi: 10.3389/fpsyg.2013.00755
- Peters, F., Villeneuve, S., and Belleville, S. (2014). Predicting progression to dementia in elderly subjects with mild cognitive impairment using both cognitive and neuroimaging predictors. *J. Alzheimers Dis.* 38, 307–318. doi: 10.3233/JAD-130842
- Ramírez, J., Górriz, J. M., Segovia, F., Chaves, R., Salas-Gonzalez, D., and López, M. (2010). Computer aided diagnosis system for the Alzheimer's disease based on partial least squares and random forest SPECT image classification. *Neurosci. Lett.* 472, 99–103. doi: 10.1016/j.neulet.2010.01.056
- Ritter, K., Schumacher, J., Weygandt, M., Buchert, R., Alfeld, C., and Haynes, J.-D. (2015). Multimodal prediction of conversion to Alzheimer's disease based on

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fnagi.2018.00135/full#supplementary-material>

- incomplete biomarkers. *Alzheimers Dement.* 1, 206–215. doi: 10.1016/j.dadm.2015.01.006
- Runtti, H., Mattila, J., Van Gils, M., Koikkalainen, J., Soininen, H., and Lötjönen, J. (2014). Quantitative evaluation of disease progression in a longitudinal mild cognitive impairment cohort. *J. Alzheimers Dis.* 39, 49–61. doi: 10.3233/JAD-130359
- Salvatore, C., Battista, P., and Castiglioni, I. (2015a). Frontiers for the early diagnosis of AD by means of MRI brain imaging and support vector machines. *Curr. Alzheimer Res.* 13, 509–533.
- Salvatore, C., Cerasa, A., Battista, P., Gilardi, M. C., Quattrone, A., and Castiglioni, I. (2015b). Magnetic resonance imaging biomarkers for the early diagnosis of Alzheimer's disease: a machine learning approach. *Front. Neurosci.* 9:307. doi: 10.3389/fnins.2015.00307
- Sharif Razavian, A., Azizpour, H., Sullivan, J., and Carlsson, S. (2014). "CNN features off-the-shelf: an astounding baseline for recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, Washington, DC. 806–813.
- Sperling, R., Mormino, E., and Johnson, K. (2014). The evolution of preclinical Alzheimer's disease: implications for prevention trials. *Neuron* 84, 608–622. doi: 10.1016/j.neuron.2014.10.038
- Sperling, R. A., Jack, C. R. Jr., Black, S. E., Frosch, M. P., Greenberg, S. M., Hyman, B. T., et al. (2011). Amyloid-related imaging abnormalities in amyloid-modifying therapeutic trials: recommendations from the Alzheimer's association research roundtable workgroup. *Alzheimers Dement.* 7, 367–385. doi: 10.1016/j.jalz.2011.05.2351
- Tosun, D., Chen, Y. F., Yu, P., Sundell, K., Suhy, J., Siemens, E., et al. (2016). Amyloid status imputed from a multimodal classifier including structural MRI distinguishes progressors from nonprogressors in a mild Alzheimer's disease clinical trial cohort. *Alzheimers Dement.* 12, 977–986. doi: 10.1016/j.jalz.2016.03.009
- Varma, S., and Simon, R. (2006). Bias in error estimation when using cross-validation for model selection. *BMC Bioinformatics* 7:91. doi: 10.1186/1471-2105-7-91
- Wechsler, D. (1987). *Manual for Wechsler Memory Scale - Revised*. San Antonio, TX: The Psychological Corporation.
- Weiner, M. W., Veitch, D. P., Aisen, P. S., Beckett, L. A., Cairns, N. J., Green, R. C., et al. (2017). Recent publications from the Alzheimer's disease neuroimaging initiative: reviewing progress toward improved AD clinical trials. *Alzheimers Dement.* 13, e1–e85. doi: 10.1016/j.jalz.2016.11.007
- Wernick, M. N., Yang, Y., Brankov, J. G., Yourganov, G., and Strother, S. C. (2010). Machine learning in medical imaging. *IEEE Signal Process. Mag.* 27, 25–38. doi: 10.1109/MSP.2010.936730
- Wold, S., Ruhe, H., Wold, H., and Dunn, W. J. (1984). The collinearity problem in linear regression: the partial least squares approach to generalized inverse. *J. Sci. Stat. Comput.* 5, 735–743. doi: 10.1137/0905052
- World Alzheimer Report (2015). *The Global Impact of Dementia, An Analysis of Prevalence, Incidence, Cost and Trends*. Available at: <http://www.alz.co.uk/research/WorldAlzheimerReport2015-sheet.pdf>
- Ye, J., Farnum, M., Yang, E., Verbeeck, R., Lobanov, V., Raghavan, N., et al. (2012). Sparse learning and stability selection for predicting MCI to AD conversion using baseline ADNI data. *BMC Neurol.* 12:46. doi: 10.1186/1471-2377-12-46

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.

Copyright © 2018 Salvatore, Cerasa and Castiglioni. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Atrophy in the Thalamus But Not Cerebellum Is Specific for C9orf72 FTD and ALS Patients – An Atlas-Based Volumetric MRI Study

Sonja Schönecker¹, Christiane Neuhofer¹, Markus Otto², Albert Ludolph², Jan Kassubek², Bernhard Landwehrmeyer², Sarah Anderl-Straub², Elisa Semler², Janine Diehl-Schmid³, Catharina Prix¹, Christian Vollmar¹, Juan Fortea⁴, Deutsches FTLD-Konsortium^{2†}, Hans-Jürgen Huppertz⁵, Thomas Arzberger⁶, Dieter Edbauer^{7,8,9}, Berend Feddersen¹⁰, Marianne Dieterich^{1,7,9}, Matthias L. Schroeter^{11,12}, Alexander E. Volk¹³, Klaus Fließbach^{14,15}, Anja Schneider^{14,15}, Johannes Kornhuber¹⁶, Manuel Maler¹⁶, Johannes Prudlo^{17,18}, Holger Jahn^{19,20}, Tobias Boeckh-Behrens²¹, Adrian Danek¹, Thomas Klopstock^{7,9,22} and Johannes Levin^{1,7*}

OPEN ACCESS

Edited by:

Javier Ramírez,
University of Granada, Spain

Reviewed by:

Stavros I. Dimitriadis,
Cardiff University, United Kingdom
Iman Beheshti,
National Center of Neurology
and Psychiatry, Japan

*Correspondence:

Johannes Levin
johannes.levin@med.uni-
muenchen.de

[†] Clinical contributions came from
members of the Deutsches
FTLD-Konsortium: Carola Roßmeier,
Franziska Albrecht, Katharina
Schümborg, Sandrine Bisenius.

Received: 29 September 2017

Accepted: 12 February 2018

Published: 15 March 2018

Citation:

Schönecker S, Neuhofer C, Otto M, Ludolph A, Kassubek J, Landwehrmeyer B, Anderl-Straub S, Semler E, Diehl-Schmid J, Prix C, Vollmar C, Fortea J, Deutsches FTLD-Konsortium, Huppertz H-J, Arzberger T, Edbauer D, Feddersen B, Dieterich M, Schroeter ML, Volk AE, Fließbach K, Schneider A, Kornhuber J, Maler M, Prudlo J, Jahn H, Boeckh-Behrens T, Danek A, Klopstock T and Levin J (2018) Atrophy in the Thalamus But Not Cerebellum Is Specific for C9orf72 FTD and ALS Patients – An Atlas-Based Volumetric MRI Study. *Front. Aging Neurosci.* 10:45. doi: 10.3389/fnagi.2018.00045

¹ Department of Neurology, Ludwig Maximilians Universität München, Munich, Germany, ² Department of Neurology, University of Ulm, Ulm, Germany, ³ Department of Psychiatry and Psychotherapy, Technical University of Munich, Munich, Germany, ⁴ Hospital San Pau Barcelona, Barcelona, Spain, ⁵ Swiss Epilepsy Clinic, Klinik Lengg, Zurich, Switzerland, ⁶ Center for Neuropathology and Prion Research, Ludwig Maximilians Universität München, Munich, Germany, ⁷ German Center for Neurodegenerative Diseases (DZNE), Munich, Germany, ⁸ Institute for Metabolic Biochemistry, Ludwig Maximilians Universität München, Munich, Germany, ⁹ Munich Cluster for Systems Neurology (SyNergy), Munich, Germany, ¹⁰ Department of Palliative Medicine, Ludwig Maximilians Universität München, Munich, Germany, ¹¹ Max Planck Institute for Human Cognitive and Brain Sciences (MPG), Leipzig, Germany, ¹² Clinic for Cognitive Neurology, University Hospital Leipzig, Leipzig, Germany, ¹³ Institute of Human Genetics, University Medical Center Hamburg-Eppendorf, Hamburg, Germany, ¹⁴ German Center for Neurodegenerative Diseases (DZNE), Bonn, Germany, ¹⁵ Department for Neurodegenerative Diseases and Geriatric Psychiatry, University Hospital Bonn, Bonn, Germany, ¹⁶ Department of Psychiatry and Psychotherapy, Friedrich-Alexander University Erlangen-Nürnberg, Erlangen, Germany, ¹⁷ Department of Neurology, Rostock University Medical Center, Rostock, Germany, ¹⁸ German Center for Neurodegenerative Diseases (DZNE), Rostock, Germany, ¹⁹ Department of Psychiatry and Psychotherapy, University Medical Center Hamburg-Eppendorf, Hamburg, Germany, ²⁰ AMEOS Klinikum Heiligenhafen, Heiligenhafen, Germany, ²¹ Department of Diagnostic and Interventional Neuroradiology, Technical University of Munich, Munich, Germany, ²² Friedrich Baur Institute at the Department of Neurology, Ludwig Maximilians Universität München, Munich, Germany

Background: The neuropathology of patients with frontotemporal dementia (FTD) or amyotrophic lateral sclerosis (ALS) due to a C9orf72 mutation is characterized by two distinct types of characteristic protein depositions containing either TDP-43 or so-called dipeptide repeat proteins that extend beyond frontal and temporal regions. Thalamus and cerebellum seem to be preferentially affected by the dipeptide repeat pathology unique to C9orf72 mutation carriers.

Objective: This study aimed to determine if mutation carriers showed an enhanced degree of thalamic and cerebellar atrophy compared to sporadic patients or healthy controls.

Methods: Atlas-based volumetry was performed in 13 affected C9orf72 FTD, ALS and FTD/ALS patients, 45 sporadic FTD and FTD/ALS patients and 19 healthy controls. Volumes and laterality indices showing significant differences between mutation carriers and sporadic patients were subjected to binary logistic regression to determine the best predictor of mutation carrier status.

Results: Compared to sporadic patients, mutation carriers showed a significant volume reduction of the thalamus, which was most striking in the occipital, temporal and

prefrontal subregion of the thalamus. Disease severity measured by mini mental status examination (MMSE) and FTD modified Clinical Dementia Rating Scale Sum of Boxes (FTD-CDR-SOB) significantly correlated with volume reduction in the aforementioned thalamic subregions. No significant atrophy of cerebellar regions could be detected. A logistic regression model using the volume of the prefrontal and the laterality index of the occipital subregion of the thalamus as predictor variables resulted in an area under the curve (AUC) of 0.88 while a model using overall thalamic volume still resulted in an AUC of 0.82.

Conclusion: Our data show that thalamic atrophy in *C9orf72* mutation carriers goes beyond the expected atrophy in the prefrontal and temporal subregion and is in good agreement with the cortical atrophy pattern described in *C9orf72* mutation carriers, indicating a retrograde degeneration of functionally connected regions. Clinical relevance of the detected thalamic atrophy is illustrated by a correlation with disease severity. Furthermore, the findings suggest MRI volumetry of the thalamus to be of high predictive value in differentiating *C9orf72* mutation carriers from patients with sporadic FTD.

Keywords: *C9orf72*, frontotemporal dementia, amyotrophic lateral sclerosis, atlas based volumetric MRI analysis, thalamus, cerebellum, salience network

INTRODUCTION

Frontotemporal dementia (FTD) and amyotrophic lateral sclerosis (ALS) are heterogeneous neurodegenerative disorders that are associated with one another in approximately 15% of the cases (Lomen-Hoerth et al., 2002). FTD can present with socially inappropriate behavior, apathy, lack of empathy, changes in diet and compulsive behaviors. Compared to Alzheimer's disease there is typically a relative preservation of memory and visuospatial function (Perry and Miller, 2013). ALS is a motor neuron disease that is characterized by progressive degeneration of upper and lower motor neurons. It typically manifests with progressive muscle weakness, muscular atrophy, spasticity and fasciculations (Lomen-Hoerth et al., 2002). The most common known cause of familial FTD, familial ALS or patients with a mixed presentation of both diseases (FTD/ALS) is a hexanucleotide expansion mutation in a non-coding region of *C9orf72* (Cruts et al., 2013). Compared to sporadic patients *c9orf72* mutation carriers have a greater frequency of psychotic symptoms like delusions, hallucinations or paranoid ideation and show more severe memory impairment (Snowden et al., 2015).

Previous neuroimaging studies on *c9orf72* mutation carriers have shown relatively symmetrical atrophy most prominent in the frontotemporal cortex and the insula, in keeping with the atrophy pattern described in sporadic patients (Boxer et al., 2011; Mahoney et al., 2012; Whitwell et al., 2012). In contrast to sporadic patients, however, *c9orf72* mutation carriers appear

to have more parietal and occipital cortical atrophy creating a more diffuse cortical atrophy pattern. Furthermore bithalamic and cerebellar involvement have been described (Yokoyama and Rosen, 2012; Bede et al., 2013; Prado et al., 2015; Floeter et al., 2016). Moreover, volumetric imaging data from the genetic frontotemporal dementia initiative (GenFI) show an early affection of thalamus and cerebellum in *C9orf72* mutation carriers compared to healthy controls (HC) as well as to *GRN* and *MAPT* mutation carriers (Rohrer et al., 2015).

A neuropathological hallmark of *C9orf72* mutation carrier status is the intracellular deposition of five dipeptide repeat proteins (DPR). These proteins are a product of repeat-associated non-ATG translation from sense and antisense transcripts (Ash et al., 2013; Gendron et al., 2013; Mori et al., 2013a,b). Most frequent dipeptide repeat pathology, particular inclusions of poly (GP) or poly (GA) can be detected in neocortex, hippocampus and cerebellum but dipeptide repeat inclusions are also abundant in thalamus (Schludi et al., 2015).

These recent neuropathological and neuroimaging findings provide evidence for an underappreciated role of the cerebellum and thalamus in the pathogenesis of FTD and ALS caused by a repeat expansion in *C9orf72* (Prado et al., 2015; Rohrer et al., 2015; Schludi et al., 2015). We hypothesized that thalamus and cerebellum show an enhanced degree of atrophy in *C9orf72* expansion carriers compared to sporadic patients and that thalamic atrophy goes beyond the expected atrophy in the prefrontal and temporal subregion of the thalamus in *C9orf72* mutation carriers. In the current study, we therefore aimed to elucidate the regional brain atrophy focusing on thalamic and cerebellar atrophy in *C9orf72* mutation carriers compared to patients with sporadic FTD or FTD/ALS and healthy controls.

Abbreviations: ALS, amyotrophic lateral sclerosis; AUC, area under the curve; DPR, dipeptide repeat proteins; FTD, frontotemporal dementia; FTD-CDR-SOB, FTD modified Clinical Dementia Rating Scale Sum of Boxes; HC, healthy controls; LI, laterality index; LPBA40, LONI Probabilistic Atlas; MDN, mediodorsal nucleus; MMSE, mini mental status examination; OTH, Oxford Thalamic Connectivity Atlas; ROC, receiver operating characteristic; SN, salience network.

MATERIALS AND METHODS

Ethics Statement

The study was performed according to the declaration of Helsinki (1991). Ethical approval for conduction of the study has been obtained at the coordinating site at the University of Ulm and all participating centers of the German consortium for frontotemporal lobar degeneration. Written informed consent was obtained from every participant.

Subjects

A total of 77 participants from the cohort of the German consortium for frontotemporal lobar degeneration (Otto et al., 2011) were included in the study: 13 symptomatic *C9orf72* mutation carriers (8 FTD, 2 ALS, 3 FTD/ALS), 45 with sporadic FTD (Medford and Critchley, 2010) or FTD/ALS (Snowden et al., 2015) in whom a pathological *C9orf72* expansion, *MAPT* or *GRN* mutation has been excluded and 19 healthy elderly control subjects. Diagnosis was made according to current international consensus criteria (Brooks et al., 2000; Rascovsky et al., 2011). Demographic features of participants are listed in **Table 1**.

Participants underwent general cognitive screening using the mini mental status examination (MMSE). To quantify the severity of dementia symptoms the FTD modified Clinical Dementia Rating Scale Sum of Boxes (FTD-CDR-SOB) (Knopman et al., 2008) score was used. Additionally age, education, disease duration and the occurrence of a positive family history were assessed.

MRI Acquisition

All patients and controls underwent whole-brain T1-weighted MRI on 3T scanners, and on 1.5T scanners, where 3T scanning was not available. An array head coil with a minimum of 8 channels was used. 3D-MPRAGE sequences were acquired in sagittal orientation with 1 mm x 1 mm in-plane resolution, slice thickness 1 mm, and TR/TE = 2300/2.03 ms.

MRI Data Processing and Volumetric Analysis

After pseudonymization and conversion from DICOM to ANALYZE 7.5 format the 3D T1-weighted images were processed by a fully automated and observer-independent method of atlas- and mask-based MRI volumetry using the Statistical Parametric Mapping 12 software (Wellcome Trust Centre for Neuroimaging, London, United Kingdom)¹. The method has been described in detail elsewhere (Huppertz et al., 2010, 2016; Opfer et al., 2016) and was already applied to neurodegenerative diseases in various cross-sectional and longitudinal studies (Kassubek et al., 2011; Frings et al., 2012, 2014; Höglinger et al., 2014; Huppertz et al., 2016; Schönecker et al., 2016). In short, each T1 image was normalized to Montreal Neurological Institute (MNI) template space using diffeomorphic anatomical registration through exponentiated Lie algebra (DARTEL) (Ashburner, 2007) and segmented into gray

matter, white matter, and cerebrospinal fluid components using the 'unified segmentation' algorithm of Statistical Parametric Mapping 12 with default parameters. The DARTEL algorithm is a highly elastic registration method resulting in a more precise registration of the individual brain to MNI space than the normalization methods in previous SPM versions, thereby also improving the adaptation to the space of the atlases used in the further post-processing. The volumes of specific brain structures and compartments were calculated by voxel-by-voxel multiplication and subsequent integration of normalized and modulated component images (gray matter, white matter or cerebrospinal fluid) with predefined masks in the same space. These masks are derived from different probabilistic brain atlases, such as the LONI Probabilistic Brain Atlas (LPBA40) provided by the Laboratory of Neuroimaging (LONI) at the University of California, Los Angeles, United States² (Shattuck et al., 2008) and the probabilistic thalamic connectivity atlas provided by the Nuffield Department of Clinical Neurosciences at the University of Oxford, United Kingdom (Oxford Thalamic Connectivity Atlas; OTH)³ (Behrens et al., 2003). Target structures were chosen *a priori* for analysis of group differences in volume (13 in total, see **Table 2**). As our study aimed to determine the amount of thalamic and cerebellar atrophy of *C9orf72* mutation carriers compared to sporadic patients and HC, we included as regions of interest cerebellum, cerebellar vermis plane and pons as derived from structures and further parcellations in the LPBA40 atlas (Huppertz et al., 2016) and in addition all structures of the OTH atlas, i.e., overall thalamic volume as well as the primary motor, sensory, posterior parietal, occipital, temporal and prefrontal subregion of the thalamus that are connected to the corresponding cortical zone. Furthermore, since frontotemporal cortex shows pronounced atrophy in *C9orf72* mutation carriers as well as in sporadic patients the frontal and temporal cortex as derived from the integration of single gyri of the LPBA40 atlas (Huppertz et al., 2010) have been included as regions of interest as well.

Statistical Analysis

Data were analyzed using SPSS23. Non-dichotomized mean scores of demographic and neuropsychological data were compared across the three groups (*C9orf72* mutation carriers, sporadic patients and HC) via Kruskal–Wallis test and Mann–Whitney test. Chi-square analysis was used to check for significant differences in gender and family history across all groups. Standard statistical significance level was set at $p < 0.05$.

For each region of interest the individual volume at clinic presentation was determined in ml. For comparison, the measured volumes were corrected by individual intracranial volume and standardized to the mean intracranial volume of healthy controls.

For group comparisons of volumetric data, a Kruskal–Wallis test was performed. Significance levels for the Kruskal–Wallis test were adjusted according to Bonferroni correction ($p < 0.0038$).

¹ www.fil.ion.ucl.ac.uk/spm

² <http://www.loni.usc.edu/atlas/>

³ <http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/Atlases>

TABLE 1 | Demographics and neuropsychological measures of the study sample.

| | <i>C9orf72</i> | Sporadic | HC | <i>C9orf72</i> vs. HC | <i>C9orf72</i> vs. sporadic | sporadic vs. HC |
|-------------------------------|----------------|------------|-------------|-----------------------|-----------------------------|-----------------|
| N | 13 | 45 | 19 | | | |
| Gender (M/F) | 8/5 | 26/19 | 12/7 | n.s. | n.s. | n.s. |
| Positive family history (Y/N) | 8/5 | 12/33 | 1/18 | + | + | n.s. |
| Age | 64.1 (8.5) | 62.8 (9.4) | 65.9 (10.1) | n.s. | n.s. | n.s. |
| Education | 12.8 (2.7) | 13.2 (3.3) | 13.7 (3.1) | n.s. | n.s. | n.s. |
| Disease duration | 2.5 (3.5) | 3.2 (3.7) | | | n.s. | |
| MMSE | 25.8 (3.1) | 25.6 (4.1) | 29.0 (0.9) | + | n.s. | + |
| FTLD-CDR-SOB | 3.7 (1.0) | 6.3 (2.9) | 0.4 (0.6) | + | n.s. | + |

n.s., not significant; + $p < 0.05$.

TABLE 2 | Anatomical structures selected for volumetric analysis per group mean and SD (in ml), and pairwise *post hoc* Bonferroni test results.

| Dependent variable | <i>C9orf72</i> | Sporadic | HC | <i>C9orf72</i> vs. HC | <i>C9orf72</i> vs. sporadic | Sporadic vs. HC |
|-------------------------|-----------------|----------------|-----------------|-----------------------|-----------------------------|-----------------|
| Frontal | 246.90 (26.96) | 265.37 (33.10) | 296.54 (15.82) | + | n.s. | + |
| Temporal | 158.31 (11.97) | 164.44 (18.78) | 181.70 (14.08) | + | n.s. | + |
| Cerebellum | 109.46 (9.82) | 109.44 (11.43) | 113.33 (10.24) | n.s. | n.s. | n.s. |
| Cerebellar vermis plane | 949.72 (101.54) | 973.68 (96.92) | 1020.74 (86.26) | n.s. | n.s. | n.s. |
| Pons | 15.30 (1.40) | 15.23 (1.76) | 15.25 (1.36) | n.s. | n.s. | n.s. |
| Thalamic regions | | | | | | |
| Thalamus all | 15.70 (1.49) | 17.62 (1.62) | 19.09 (1.50) | + | + | + |
| Primary motor | 0.97 (0.08) | 1.03 (0.07) | 1.06 (0.08) | n.s. | n.s. | n.s. |
| Sensory | 1.11 (0.09) | 1.18 (0.09) | 1.23 (0.09) | + | + | n.s. |
| Premotor | 1.70 (0.15) | 1.84 (0.14) | 1.92 (0.14) | + | + | n.s. |
| Posterior parietal | 3.09 (0.27) | 3.35 (0.26) | 3.50 (0.25) | + | + | n.s. |
| Occipital | 1.41 (0.13) | 1.62 (0.18) | 1.77 (0.16) | + | + | + |
| Temporal | 2.91 (0.29) | 3.45 (0.48) | 3.92 (0.43) | + | + | + |
| Prefrontal | 4.52 (0.37) | 5.14 (0.51) | 5.70 (0.46) | + | + | + |

n.s., not significant; + $p < 0.05$, corrected for multiple comparisons.

Results of post-hoc tests were regarded significant if they survived an additional Bonferroni correction for multiple pairwise group comparisons. Spearman's test was used to explore significant correlations between volumetric data and neuropsychological variables. Significance level was adjusted according to Bonferroni correction ($p < 0.0038$) as well.

Furthermore, to assess laterality of overall thalamic volume and thalamic subregions a laterality index (LI) defined as the ratio $[(\text{left} - \text{right})/(\text{left} + \text{right})]$ was calculated (Seghier, 2008; Okada et al., 2016) for each study group. LIs can range from -1 to 1 with a positive LI indicating a leftward asymmetry. One-sample Wilcoxon signed rank tests were calculated to evaluate whether LIs were significantly different from zero. A Kruskal-Wallis test was performed for group comparisons of LIs. As for the volumetric analysis significance level was adjusted according to Bonferroni correction ($p < 0.00625$) and an additional *post hoc* Bonferroni correction was performed.

Parameters showing a significant difference between *C9orf72* mutation carriers and sporadic patients in the former analyses, i.e., overall thalamic volume and the volumes of the sensory, premotor, parietal, occipital, temporal and prefrontal subregion of the thalamus, the LI of the primary motor, sensory, premotor, occipital and prefrontal subregion of the thalamus as well as the neuropsychological parameters MMSE and FTD-CDR-SOB

were subjected to a forward stepwise binary logistic regression to determine the best predictor of diagnosis. Furthermore, the receiver operating characteristic (ROC) curve was created to evaluate the utility of the model at distinguishing between *C9orf72* mutation carriers and sporadic patients.

RESULTS

Demographics and Cognitive Scores

Demographics and cognitive scores of the study sample can be seen in **Table 1**. Participant groups did not differ in terms of gender, age, education and disease duration. Patient groups (*C9orf72* mutation carriers and sporadic patients) performed significantly worse at cognitive screening tests (MMSE and FTD-CDR-SOB) compared to HC but did not differ significantly from one another. *C9orf72* mutation carriers had, as expected, significantly more frequently a positive family history compared to sporadic patients and HC (see **Table 1**).

Volumetric Analysis

Kruskal-Wallis test revealed significant differences of the volumes of the frontal and temporal lobe, overall thalamic volume (**Figure 1**) and the volumes of the sensory, premotor,

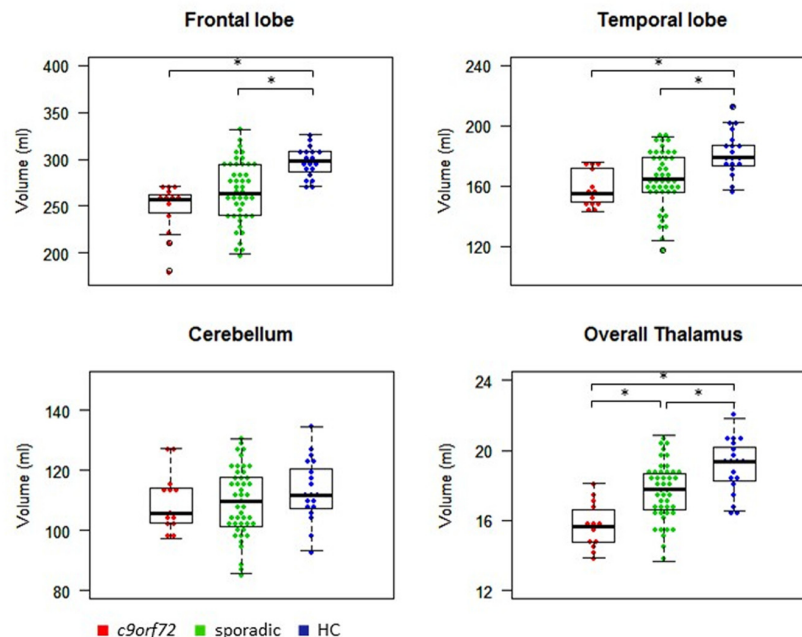


FIGURE 1 | Volumes of frontal and temporal lobe, cerebellum and thalamus. *Indicates significant differences.

posterior parietal, occipital, temporal and prefrontal subregion of the thalamus (**Figure 2**). No significant differences could be detected for the cerebellum, the cerebellar vermis plane and pons as well as for the primary motor subregion of the thalamus (**Figures 1, 2**).

Post hoc Bonferroni tests showed that sporadic patients compared to HC had reduced volumes of the frontal and temporal lobe as well as reduced volumes of overall thalamus and the occipital, temporal and prefrontal subregion of the thalamus (**Figures 1, 2** and **Table 2**). Similarly, *C9orf72* mutation carriers had significantly smaller volumes of frontal and temporal lobe as well as all investigated thalamic subregions apart from the primary motor subregion (i.e., sensory, premotor, posterior parietal, occipital, temporal and prefrontal subregion of the thalamus) compared to HC. No significant differences of frontal and temporal lobe volumes could be detected between *C9orf72* mutation carriers and sporadic patients. However, although sporadic patients were somewhat more advanced, both in disease duration and in the FTD-CDR-SOB, *C9orf72* mutation carriers showed significantly smaller volumes of the sensory, premotor, posterior parietal, occipital, temporal and prefrontal subregion of the thalamus (**Figure 2** and **Table 2**).

Correlation Analysis

Correlation analysis showed a significant positive/negative correlation of MMSE/FTD-CDR-SOB with overall thalamic volume ($r_s = 0.352$, $r_s = -0.406$), the volumes of the prefrontal ($r_s = 0.368$, $r_s = -0.453$), temporal ($r_s = 0.377$, $r_s = -0.423$) and occipital ($r_s = 0.367$, $r_s = -0.385$) subregion of the thalamus as well as with frontal ($r_s = 0.392$, $r_s = -0.607$) and temporal ($r_s = 0.364$, $r_s = -0.391$) lobe volume (**Figure 3**).

Laterality Indices

Laterality indices were investigated in this study as well. All study groups showed positive LIs of the occipital, prefrontal, and posterior parietal subregion of the thalamus and significantly negative LIs of the primary motor, sensory, premotor and temporal subregion of the thalamus. Overall thalamic volume showed significantly positive LIs in sporadic patients and HC but was not significantly different from zero in *C9orf72* mutation carriers.

Kruskal–Wallis test revealed significant group differences of the LIs of overall thalamus and of the primary motor, sensory, premotor, occipital and prefrontal subregion of the thalamus. Only for the posterior parietal and the temporal subregion of the thalamus, no significant group differences could be detected. *C9orf72* mutation carriers differed significantly from sporadic patients in the aforementioned LIs. Furthermore LIs of overall thalamus, sensory, premotor, occipital and prefrontal subregion of the thalamus differed significantly between *C9orf72* mutation carriers and HC. No significant differences of LIs could be detected between sporadic patients and HC. For every investigated thalamic subregion as well as for overall thalamus, *C9orf72* mutation carriers showed the lowest LIs (**Table 3**).

Binary Logistic Regression

Target volumes, LIs and neuropsychological variables that showed a significant difference between *C9orf72* mutation carriers and sporadic patients were subjected to a forward stepwise binary logistic regression to determine the best predictor of diagnosis. The optimal logistic regression model using the volume of the prefrontal subregion of the thalamus ($B = -2.54$) and the LI of the occipital subregion of the thalamus ($B = -24.81$)

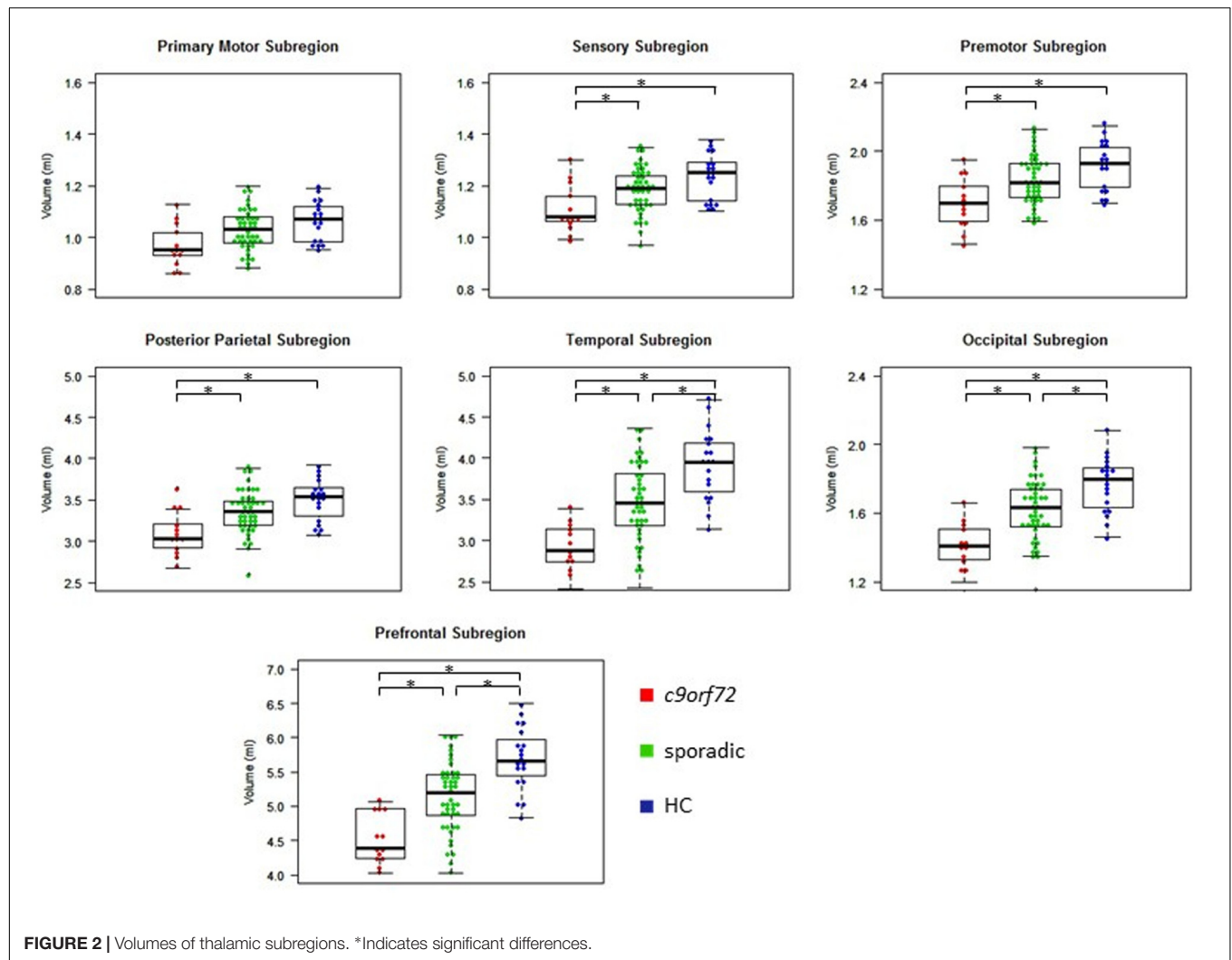


FIGURE 2 | Volumes of thalamic subregions. *Indicates significant differences.

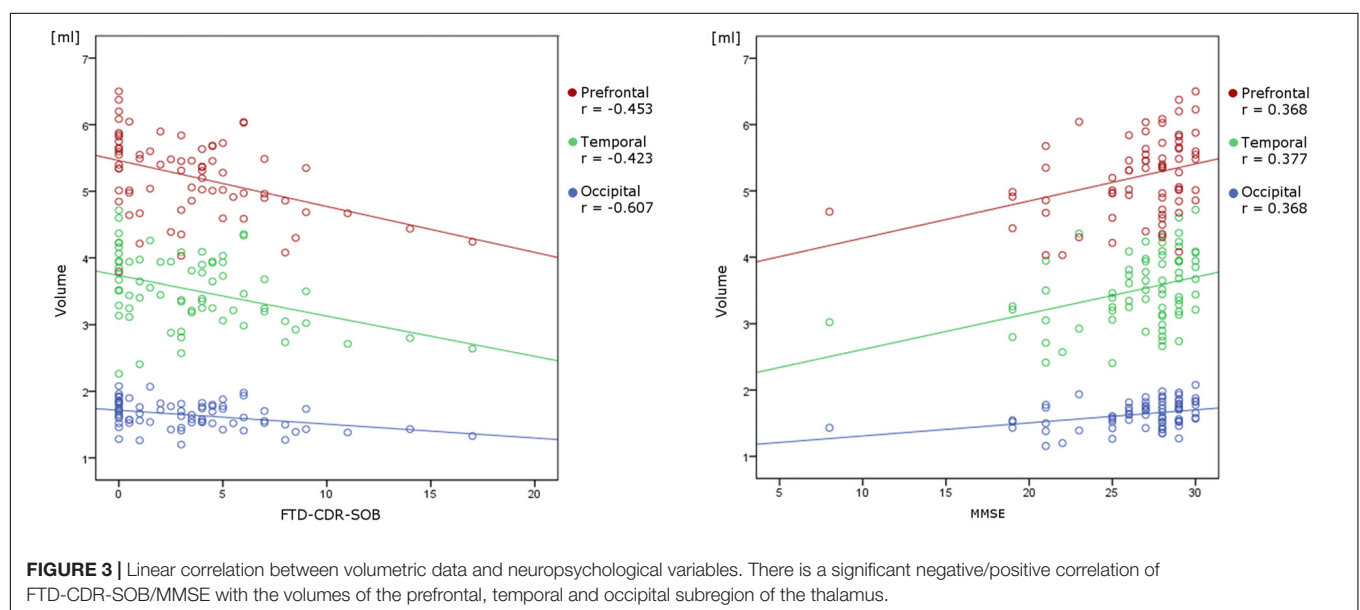


FIGURE 3 | Linear correlation between volumetric data and neuropsychological variables. There is a significant negative/positive correlation of FTD-CDR-SOB/MMSE with the volumes of the prefrontal, temporal and occipital subregion of the thalamus.

TABLE 3 | Laterality indices of overall thalamus and thalamic subregions, and pairwise *post hoc* Bonferroni test results.

| Dependent variable | <i>C9orf72</i> | Sporadic | HC | <i>C9orf72</i> vs. HC | <i>C9orf72</i> vs. sporadic | Sporadic vs. HC |
|--------------------|----------------|----------|---------|-----------------------|-----------------------------|-----------------|
| Overall thalamus | 0.0061 | 0.0327 | 0.0349 | + | + | n.s. |
| Primary motor | −0.0956 | −0.0687 | −0.0722 | n.s. | + | n.s. |
| Sensory | −0.0731 | −0.0494 | −0.0503 | + | + | n.s. |
| Premotor | −0.0795 | −0.0531 | −0.0548 | + | + | n.s. |
| Posterior parietal | 0.0360 | 0.0576 | 0.0525 | n.s. | n.s. | n.s. |
| Occipital | 0.2373 | 0.2698 | 0.2814 | + | + | n.s. |
| Temporal | −0.0713 | −0.0448 | −0.0419 | n.s. | n.s. | n.s. |
| Prefrontal | 0.0366 | 0.0632 | 0.0690 | + | + | n.s. |

n.s., not significant; + $p < 0.05$, corrected for multiple comparisons.

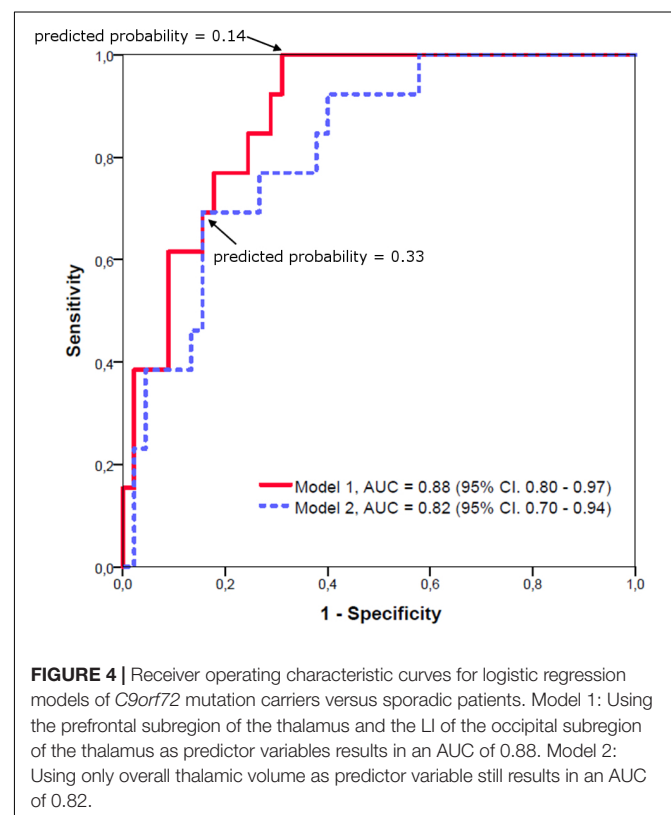
as predictor variables ($p < 0.05$; Nagelkerke's $R^2 = 0.441$) resulted in an area under the curve (AUC) of 0.88 (95% CI: 0.80 – 0.97). The highest combination of sensitivity and specificity was acquired with a sensitivity of 1.00 and a specificity of 0.69 at a predicted probability cutoff of 0.14.

A logistic regression model using only overall thalamic volume as predictor variable ($B = -0.82$, $p < 0.05$; Nagelkerke's $R^2 = 0.327$) still resulted in an AUC of 0.82 (95% CI: 0.70 – 0.94). The highest combination of sensitivity and specificity was acquired with a sensitivity of 0.69 and a specificity of 0.84 at a predicted probability cutoff of 0.33. The ROC curves are shown in **Figure 4**.

DISCUSSION

Neuropathological data show that DPR protein aggregates are abundant in the thalamus (Schludi et al., 2015). This is in accordance with previous reports of significant thalamic atrophy in *C9orf72* mutation carriers (Sha et al., 2012; Mahoney et al., 2012; Irwin et al., 2013). Furthermore, an early affection of the thalamus in presymptomatic *C9orf72* mutation carriers has been shown (Rohrer et al., 2015). Here we show a significant volume reduction of certain thalamic subregions of symptomatic *C9orf72* mutation carriers compared to sporadic patients and HC and reveals overall thalamic volume to be a useful predictor of *C9orf72* mutation carrier status. The negative correlation of thalamic volume and clinical parameters highlights the important role of the thalamus in the pathogenesis of *C9orf72* associated clinical pictures of FTD and ALS.

C9orf72 mutation carriers present more frequently with psychotic symptoms and show more severe memory impairment than sporadic patients (Boeve et al., 2012; Snowden et al., 2015). In our study cohort, disease severity measured by MMSE and FTD-CDR-SOB was not only correlated with the volumes of the frontal and temporal lobes but also with overall thalamic volume and the volumes of the prefrontal, temporal and occipital subregion of the thalamus, illustrating their clinical relevance. These subregions furthermore showed the most striking volumetric differences between *C9orf72* mutation carriers and sporadic patients. The occipital subregion includes lateral geniculate nucleus and parts of the inferior pulvinar, temporal subregion includes parts of the mediodorsal nucleus (MDN) and



the medial and inferior pulvinar and the prefrontal subregion includes some of MDN, ventral anterior nucleus and anterodorsal and anteromedial nucleus (Behrens et al., 2003; Johansen-Berg et al., 2005).

The “Salience Network (SN)” (Seeley et al., 2007) is an intrinsic connectivity network that is activated in response to emotionally significant stimuli and contributes to complex brain functions like guidance of behavioral responses, production of subjective feelings and initiation of cognitive control (Menon and Uddin, 2010; Medford and Critchley, 2010). The SN is anchored in the dorsal anterior cingulate and anterior insula but also includes various subcortical structures. Comparable SN disruption despite contrasting atrophy patterns in *C9orf72* mutation carriers and sporadic patients has been described (Lee et al., 2014). Atrophy of the medial pulvinar nucleus that has prominent reciprocal

connections to major hubs of the SN (Romanski et al., 1997) could only be detected in *C9orf72* mutation carriers. As medial pulvinar nucleus atrophy predicted reduced SN connectivity, a strategic atrophy of the medial pulvinar nucleus has been proposed to contribute to SN disruption in *C9orf72* mutation carriers (Lee et al., 2014). We therefore hypothesize that especially atrophy of the medial pulvinar nucleus may have led to the detected volume reduction of the temporal subregion of the thalamus of *C9orf72* mutation carriers in our study cohort.

Another thalamic node that is part of the SN is the MDN. Studies detected significant atrophy of MDN in early stages of FTD (Seeley et al., 2008). MDN is believed to be involved in the pathogenesis of schizophrenia (Young et al., 2000; Alelú-Paz and Giménez-Amaya, 2008) and has been shown to play an important role in working memory and episodic memory (Gaffan and Parker, 2000; Watanabe and Funahashi, 2012; Mitchell and Chakraborty, 2013). It is involved in memory formation and influences emotional connotations also via extralemniscal pathways, e.g., connecting to the amygdala. For ALS a more frequent bulbar onset in mutation carriers is discussed. With view to FTD, mutation carriers may present with more psychosis displaying delusions and hallucinations and/or catatonic features. Also, late-onset dementia and depressive syndromes with cognitive impairments were reported (Ducharme et al., 2017). We hypothesize that MDN atrophy is reflected by the volume reduction of the temporal and prefrontal subregion of the thalamus of *C9orf72* mutation carriers and that MDN atrophy leads to a disruption of SN and thereby contributes to the distinct clinical characteristics of mutation and non-mutations carriers.

Several studies have reported greater occipital (Boxer et al., 2011; Khan et al., 2012; Whitwell et al., 2012) and parietal (Sha et al., 2012; Whitwell et al., 2012) volume loss of *C9orf72* mutation carriers compared to sporadic patients. The detected volume loss of the occipital, posterior parietal and sensory subregion of the thalamus of *C9orf72* mutation carriers may therefore be due to a common degeneration of functionally connected regions. The fact that the atrophy pattern in our *C9orf72* mutation carrier study group goes beyond the expected atrophy in the frontal and temporal subregion of the thalamus is an indicator that atrophy in mutation carriers may exceed the SN and is in keeping with the described cortical pattern of atrophy detected in mutation carriers that also goes beyond the sporadic FTD-associated atrophy pattern.

In a recent volumetric MRI study, a classification accuracy of 93% could be obtained to discriminate between *C9orf72* mutation carriers, *MAPT* and *GRN* mutation carriers and sporadic patients by using 26 regional volume and asymmetry scores (Whitwell et al., 2012). A more conservative model requiring 14 variables was able to classify 74% of patients correctly. In contrast to the aforementioned study, we compared only *C9orf72* mutation carriers and sporadic patients. As group sizes differed in our study cohort, classification accuracy cannot directly be compared. Furthermore, as multicollinearity was present in our data, results of binary logistic regression have to be interpreted with caution. However, multicollinearity does

not bias the result of logistic regression, but only affects calculations regarding individual predictor variables (Midi et al., 2010). The optimal logistic regression model resulted in an AUC of 0.88 while a logistic regression model using only overall thalamic volume as predictor variable still resulted in an AUC of 0.82. Both AUCs correspond to very good diagnostic accuracy (Šimundić, 2008). Our data therefore provide evidence of a combination of the volume of the prefrontal subregion and the LI of the occipital subregion of the thalamus and overall thalamic volume respectively to be of high predictive value in identifying *C9orf72* mutation carriers. MRI volumetry, especially of subcortical regions of interest, may therefore help to differentiate between *C9orf72* mutation carriers and sporadic patients, regardless of the presence of a positive family history. This is particularly useful since prediction of mutation status is extremely difficult based on clinical features alone. Nonetheless, the addition of clinical information like prominent psychosis or memory impairment (Boeve et al., 2012; Snowden et al., 2015), co-occurring FTD and ALS symptoms (DeJesus-Hernandez et al., 2011) and a positive family history may further improve prediction.

As in previous studies examining thalamic asymmetries in control subjects, we were able to detect left greater than right asymmetry in our HC (Flaum et al., 1995; Okada et al., 2016). Leftward asymmetry could also be detected in sporadic patients. Mainly the posterior parietal, occipital and prefrontal subregion of the thalamus seem to contribute to the detected overall leftward asymmetry, whereas the primary motor, premotor, sensory and temporal subregion of the thalamus display a rightward asymmetry. In contrast, although each thalamic subregion showed either rightward or leftward asymmetry, no significant overall thalamic asymmetry could be detected in *C9orf72* mutation carriers. This is consistent with the symmetric cortical atrophy pattern detected in *C9orf72* mutation carriers (Mahoney et al., 2012; Whitwell et al., 2012).

Although abundant DPR pathology within granule cell layer of the cerebellum seems to be a consistent finding in *C9orf72* mutation carriers (Al-Sarraj et al., 2011; Irwin et al., 2013) and a number of studies have reported more prominent cerebellar atrophy in *C9orf72* mutation carriers compared to sporadic patients (Mahoney et al., 2012; Whitwell et al., 2012; Irwin et al., 2013), we were not able to detect significant group differences with respect to cerebellar volume. In a recent study, focal atrophy localized to lobule VIIa-Crus I in the superior-posterior region of the cerebellum could be detected in *C9orf72* mutation carriers compared to HC (Bocchetta et al., 2016). As this area is connected via the thalamus to the prefrontal cortex (Krienen and Buckner, 2009; Stoodley and Schmahmann, 2010) and therefore with the SN (Caulfield et al., 2016) and is associated with goal-directed behaviors, its involvement in *C9orf72*-associated FTD and ALS seems plausible. Perhaps an investigation of cerebellar subregions would have revealed atrophy clusters specific for *C9orf72* mutation carriers.

A limitation of the current study that needs to be considered is the small number of *C9orf72* mutation carriers enrolled ($N = 13$) which rendered subdividing *C9orf72* mutation carriers in FTD, FTD/ALS and ALS patients impossible. Further studies

of larger cohorts subdividing the different *C9orf72* mutation carrier phenotypes are necessary. Furthermore multi-scanner data-sets and scans performed on 3T scanners and on 1.5T scanners have been pooled in the analyses. However multi-site studies offer a good possibility to investigate rare disorders like neurodegenerative diseases caused by *c9orf72* mutation carrier status. Another limitation of our study was the absence of a neuropathological confirmation of our sporadic patient study group which leaves the possibility that a percentage of cases had a mismatch of clinical diagnosis and underlying pathology.

Keeping these limitations in mind, our findings reveal a combination of the volume of the prefrontal subregion and the LI of the occipital subregion of the thalamus and overall thalamic volume respectively to be useful predictors of mutation carrier status. We furthermore demonstrated that the thalamic atrophy pattern in *C9orf72* mutation carriers goes beyond hubs of the SN and is in good agreement with the cortical atrophy pattern detected in *C9orf72* mutation carriers, indicating a retrograde degeneration of functionally connected regions.

AUTHOR CONTRIBUTIONS

SS and JL coordinated and drafted the manuscript. CN, JD-S, MO, AL, JaK, BL, SA-S, ES, KF, AS, JoK, MM, JP, HJ, TB-B, MS, and CP were involved in the imaging acquisition. AV

performed the genetic testing. SS, CV, and H-JH were involved in the data analysis. SS, JF, TA, DE, BF, MD, JoK, MS, AD, TK, and JL were involved in the interpretation of data. All authors critically revised the manuscript and read and approved the final manuscript.

FUNDING

This research was supported by the Federal Ministry of Education and Research (BMBF) by a grant given to German FTLD consortium (Grant No. O1Gl1007A) and the Munich Cluster for Systems Neurology (SyNergy), the European Community's Health Seventh Framework Program under grant agreement 617198 [DPR-MODELS] and the Lüneburg Foundation. Moreover, MS has been supported by the Michael J. Fox Foundation (MS; Grant No. 11362). The work of AV was funded by the Deutsche Forschungsgemeinschaft [DFG, VO 2028(1-1)].

ACKNOWLEDGMENTS

We would like to express our thanks to the patients and their caregivers.

REFERENCES

- Alelú-Paz, R., and Giménez-Amaya, J. M. (2008). The mediodorsal thalamic nucleus and schizophrenia. *J. Psychiatry Neurosci.* 33, 489–498.
- Al-Sarraj, S., King, A., Troakes, C., Smith, B., Maekawa, S., Bodi, I., et al. (2011). p62 positive, TDP-43 negative, neuronal cytoplasmic and intranuclear inclusions in the cerebellum and hippocampus define the pathology of *C9orf72*-linked FTL and MND/ALS. *Acta Neuropathol.* 122, 691–702. doi: 10.1007/s00401-011-0911-2
- Ash, P. E., Bieniek, K. F., Gendron, T. F., Caulfield, T., Lin, W. L., DeJesus-Hernandez, M., et al. (2013). Unconventional translation of C9ORF72 GGGGCC expansion generates insoluble polypeptides specific to c9FTD/ALS. *Neuron* 77, 639–646. doi: 10.1016/j.neuron.2013.02.004
- Ashburner, J. (2007). A fast diffeomorphic image registration algorithm. *Neuroimage* 38, 95–113. doi: 10.1016/j.neuroimage.2007.07.007
- Bede, P., Bokde, A. L., Byrne, S., Elamin, M., McLaughlin, R. L., Kenna, K., et al. (2013). Multiparametric MRI study of ALS stratified for the *C9orf72* genotype. *Neurology* 81, 361–369. doi: 10.1212/WNL.0b013e31829c5eee
- Behrens, T. E., Johansen-Berg, H., Woolrich, M. W., Smith, S. M., Wheeler-Kingshott, C. A., Boulby, P. A., et al. (2003). Non-invasive mapping of connections between human thalamus and cortex using diffusion imaging. *Nat. Neurosci.* 6, 750–757. doi: 10.1038/nn1075
- Bocchetta, M., Cardoso, M. J., Cash, D. M., Ourselin, S., Warren, J. D., and Rohrer, J. D. (2016). Patterns of regional cerebellar atrophy in genetic frontotemporal dementia. *Neuroimage Clin.* 11, 287–290. doi: 10.1016/j.nicl.2016.02.008
- Boeve, B. F., Boylan, K. B., Graff-Radford, N. R., DeJesus-Hernandez, M., Knopman, D. S., Pedraza, O., et al. (2012). Characterization of frontotemporal dementia and/or amyotrophic lateral sclerosis associated with the GGGGCC repeat expansion in C9ORF72. *Brain* 135, 765–783. doi: 10.1093/brain/aww004
- Boxer, A. L., Mackenzie, I. R., Boeve, B. F., Baker, M., Seeley, W. W., Crook, R., et al. (2011). Clinical, neuroimaging and neuropathological features of a new chromosome 9p-linked FTD-ALS family. *J. Neurol. Neurosurg. Psychiatry* 82, 196–203. doi: 10.1136/jnnp.2009.204081
- Brooks, B. R., Miller, R. G., Swash, M., Munsat, T. L., and World Federation of Neurology Research Group on Motor Neuron Diseases (2000). El Escorial revisited: revised criteria for the diagnosis of amyotrophic lateral sclerosis. *Amyotroph. Lateral Scler. Other Motor Neuron Disord.* 1, 293–299. doi: 10.1080/146608200300079536
- Caulfield, M. D., Zhu, D. C., McAuley, J. D., and Servatius, R. J. (2016). Individual differences in resting-state functional connectivity with the executive network: support for a cerebellar role in anxiety vulnerability. *Brain Struct. Funct.* 221, 3081–3093. doi: 10.1007/s00429-015-1088-6
- Cruts, M., Gijselinck, I., Van Langenhove, T., van der Zee, J., and Van Broeckhoven, C. (2013). Current insights into the *C9orf72* repeat expansion diseases of the FTL/ALS spectrum. *Trends Neurosci.* 36, 450–459. doi: 10.1016/j.tins.2013.04.010
- DeJesus-Hernandez, M., Mackenzie, I. R., Boeve, B. F., Boxer, A. L., Baker, M., Rutherford, N. J., et al. (2011). Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS. *Neuron* 72, 245–256. doi: 10.1016/j.neuron.2011.09.011
- Ducharme, S., Bajestan, S., Dickerson, B. C., and Voon, V. (2017). Psychiatric presentations of *C9orf72* mutation: What are the diagnostic implications for clinicians? *J. Neuropsychiatry Clin. Neurosci.* 29, 195–205. doi: 10.1176/appi.neuropsych.16090168
- Flaum, M., Swayze, V. W., O'Leary, D. S., Yuh, W. T., Ehrhardt, J. C., Arndt, S. V., et al. (1995). Effects of diagnosis, laterality, and gender on brain morphology in schizophrenia. *Am. J. Psychiatry* 152, 704–714. doi: 10.1176/ajp.152.5.704
- Floeter, M. K., Bageac, D., Danielian, L. E., Braun, L. E., Traynor, B. J., and Kwan, J. Y. (2016). Longitudinal imaging in *C9orf72* mutation carriers: relationship to phenotype. *Neuroimage Clin.* 12, 1035–1043. doi: 10.1016/j.nicl.2016.10.014
- Frings, L., Mader, I., Landwehrmeyer, B. G., Weiller, C., Hüll, M., and Huppertz, H. J. (2012). Quantifying change in individual subjects affected by frontotemporal lobar degeneration using automated longitudinal MRI volumetry. *Hum. Brain Mapp.* 33, 1526–1535. doi: 10.1002/hbm.21304

- Frings, L., Yew, B., Flanagan, E., Lam, B. Y., Hüll, M., Huppertz, H. J., et al. (2014). Longitudinal grey and white matter changes in frontotemporal dementia and Alzheimer's disease. *PLoS One* 9:e90814. doi: 10.1371/journal.pone.0090814
- Gaffan, D., and Parker, A. (2000). Mediodorsal thalamic function in scene memory in rhesus monkeys. *Brain* 123, 816–827. doi: 10.1093/brain/123.4.816
- Gendron, T. F., Bieniek, K. F., Zhang, Y. J., Jansen-West, K., Ash, P. E., Caulfield, T., et al. (2013). Antisense transcripts of the expanded C9ORF72 hexanucleotide repeat form nuclear RNA foci and undergo repeat-associated non-ATG translation in c9FTD/ALS. *Acta Neuropathol.* 126, 829–844. doi: 10.1007/s00401-013-1192-8
- Höglinger, G. U., Huppertz, H. J., Wagenpfeil, S., Andrés, M. V., Belloch, V., León, T., et al. (2014). Tideglusib reduces progression of brain atrophy in progressive supranuclear palsy in a randomized trial. *Mov. Disord.* 29, 479–487. doi: 10.1002/mds.25815
- Huppertz, H. J., Kröll-Seger, J., Klöppel, S., Ganz, R. E., and Kassubek, J. (2010). Intra- and interscanner variability of automated voxel-based volumetry based on a 3D probabilistic atlas of human cerebral structures. *Neuroimage* 49, 2216–2224. doi: 10.1016/j.neuroimage.2009.10.066
- Huppertz, H. J., Möller, L., Südmeyer, M., Hilker, R., Hattingen, E., Egger, K., et al. (2016). Differentiation of neurodegenerative parkinsonian syndromes by volumetric magnetic resonance imaging analysis and support vector machine classification. *Mov. Disord.* 31, 1506–1517. doi: 10.1002/mds.26715
- Irwin, D. J., McMillan, C. T., Brettschneider, J., Libon, D. J., Powers, J., Rascovsky, K., et al. (2013). Cognitive decline and reduced survival in C9orf72 expansion frontotemporal degeneration and amyotrophic lateral sclerosis. *J. Neurol. Neurosurg. Psychiatry* 84, 163–169. doi: 10.1136/jnnp-2012-303507
- Johansen-Berg, H., Behrens, T. E., Sillery, E., Ciccarelli, O., Thompson, A. J., Smith, S. M., et al. (2005). Functional-anatomical validation and individual variation of diffusion tractography-based segmentation of the human thalamus. *Cereb. Cortex* 15, 31–39. doi: 10.1093/cercor/bhh105
- Kassubek, J., Pinkhardt, E. H., Dietmaier, A., Ludolph, A. C., Landwehrmeyer, G. B., and Huppertz, H. J. (2011). Fully automated atlas-based MR imaging volumetry in Huntington disease, compared with manual volumetry. *AJNR Am. J. Neuroradiol.* 32, 1328–1332. doi: 10.3174/ajnr.A2514
- Khan, B. K., Yokoyama, J. S., Takada, L. T., Sha, S. J., Rutherford, N. J., Fong, J. C., et al. (2012). Atypical, slowly progressive behavioural variant frontotemporal dementia associated with C9ORF72 hexanucleotide expansion. *J. Neurol. Neurosurg. Psychiatry* 83, 358–364. doi: 10.1136/jnnp-2011-301883
- Knopman, D. S., Kramer, J. H., Boeve, B. F., Caselli, R. J., Graff-Radford, N. R., Mendez, M. F., et al. (2008). Development of methodology for conducting clinical trials in frontotemporal lobar degeneration. *Brain* 131(Pt 11), 2957–2968. doi: 10.1093/brain/awn234
- Krienen, F. M., and Buckner, R. L. (2009). Segregated fronto-cerebellar circuits revealed by intrinsic functional connectivity. *Cereb. Cortex* 19, 2485–2497. doi: 10.1093/cercor/bhp135
- Lee, S. E., Khazenzon, A. M., Trujillo, A. J., Guo, C. C., Yokoyama, J. S., Sha, S. J., et al. (2014). Altered network connectivity in frontotemporal dementia with C9orf72 hexanucleotide repeat expansion. *Brain* 137, 3047–3060. doi: 10.1093/brain/awu248
- Lomen-Hoerth, C., Anderson, T., and Miller, B. (2002). The overlap of amyotrophic lateral sclerosis and frontotemporal dementia. *Neurology* 59, 1077–1079. doi: 10.1212/WNL.59.7.1077
- Mahoney, C. J., Beck, J., Rohrer, J. D., Lashley, T., Mok, K., Shakespeare, T., et al. (2012). Frontotemporal dementia with the C9ORF72 hexanucleotide repeat expansion: clinical, neuroanatomical and neuropathological features. *Brain* 135, 736–750. doi: 10.1093/brain/awr361
- Medford, N., and Critchley, H. D. (2010). Conjoint activity of anterior insular and anterior cingulate cortex: awareness and response. *Brain Struct. Funct.* 214, 535–549. doi: 10.1007/s00429-010-0265-x
- Menon, V., and Uddin, L. Q. (2010). Saliency, switching, attention and control: a network model of insula function. *Brain Struct. Funct.* 214, 655–667. doi: 10.1007/s00429-010-0262-0
- Midi, H., Sarkar, S. K., and Rana, S. (2010). Collinearity diagnostics of binary logistic regression model. *J. Interdiscip. Math.* 13, 253–267. doi: 10.1080/09720502.2010.10700699
- Mitchell, A. S., and Chakraborty, S. (2013). What does the mediodorsal thalamus do? *Front. Syst. Neurosci.* 7:37. doi: 10.3389/fnsys.2013.00037
- Mori, K., Arzberger, T., Grasser, F. A., Gijssels, I., May, S., Rentzsch, K., et al. (2013a). Bidirectional transcripts of the expanded C9orf72 hexanucleotide repeat are translated into aggregating dipeptide repeat proteins. *Acta Neuropathol.* 126, 881–893. doi: 10.1007/s00401-013-1189-3
- Mori, K., Weng, S. M., Arzberger, T., May, S., Rentzsch, K., Kremmer, E., et al. (2013b). The C9orf72 GGGGCC repeat is translated into aggregating dipeptide-repeat proteins in FTD/ALS. *Science* 339, 1335–1338. doi: 10.1126/science.1232927
- Okada, N., Fukunaga, M., Yamashita, F., Koshiyama, D., Yamamori, H., Ohi, K., et al. (2016). Abnormal asymmetries in subcortical brain volume in schizophrenia. *Mol. Psychiatry* 21, 1460–1466. doi: 10.1038/mp.2015.209
- Opfer, R., Suppa, P., Kepp, T., Spies, L., Schippling, S., Huppertz, H. J., et al. (2016). Atlas based brain volumetry: How to distinguish regional volume changes due to biological or physiological effects from inherent noise of the methodology. *Magn. Reson. Imaging* 34, 455–461. doi: 10.1016/j.mri.2015.12.031
- Otto, M., Ludolph, A. C., Landwehrmeyer, B., Förstl, H., Diehl-Schmid, J., Neumann, M., et al. (2011). German consortium for frontotemporal lobar degeneration. *Nervenarzt* 82, 1002–1005. doi: 10.1007/s00115-011-3261-3
- Perry, D. C., and Miller, B. L. (2013). Frontotemporal dementia. *Semin. Neurol.* 33, 336–341. doi: 10.1055/s-0033-1359316
- Prado, L. G. R., Bicalho, I. C. S., Magalhães, D., Caramelli, P., Teixeira, A. L., and de Souza, L. C. (2015). C9ORF72 and the FTD-ALS spectrum: a systematic review of neuroimaging studies. *Dement. Neuropsychol.* 9, 413–421. doi: 10.1590/1980-57642015DN94000413
- Rascovsky, K., Hodges, J. R., Knopman, D., Mendez, M. F., Kramer, J. H., Neuhaus, J., et al. (2011). Sensitivity of revised diagnostic criteria for the behavioural variant of frontotemporal dementia. *Brain* 134, 2456–2477. doi: 10.1093/brain/awr179
- Rohrer, J. D., Nicholas, J. M., Cash, D. M., van Swieten, J., Doppler, E., Jiskoot, L., et al. (2015). Presymptomatic cognitive and neuroanatomical changes in genetic frontotemporal dementia in the genetic frontotemporal dementia initiative (GENFI) study: a cross-sectional analysis. *Lancet Neurol.* 14, 253–262. doi: 10.1016/S1474-4422(14)70324-2
- Romanski, L. M., Giguere, M., Bates, J. F., and Goldman-Rakic, P. S. (1997). Topographic organization of medial pulvinar connections with the prefrontal cortex in the rhesus monkey. *J. Comp. Neurol.* 379, 313–332. doi: 10.1002/(SICI)1096-9861(19970317)379:3<313::AID-CNE1>3.0.CO;2-6
- Schludi, M. H., May, S., Grasser, F. A., Rentzsch, K., Kremmer, E., Küpper, C., et al. (2015). Distribution of dipeptide repeat proteins in cellular models and C9orf72 mutation cases suggests link to transcriptional silencing. *Acta Neuropathol.* 130, 537–555. doi: 10.1007/s00401-015-1450-z
- Schönecker, S., Brendel, M., Huber, M., Vollmar, C., Huppertz, H. J., Teipel, S., et al. (2016). Applied multimodal diagnostics in a case of presenile dementia. *BMC Neurol.* 16:131. doi: 10.1186/s12883-016-0647-7
- Seeley, W. W., Crawford, R., Rascovsky, K., Kramer, J. H., Weiner, M., Miller, B. L., et al. (2008). Frontal paralimbic network atrophy in very mild behavioral variant frontotemporal dementia. *Arch. Neurol.* 65, 249–255. doi: 10.1001/archneurol.2007.38
- Seeley, W. W., Menon, V., Schatzberg, A. F., Keller, J., Glover, G. H., Kenna, H., et al. (2007). Dissociable intrinsic connectivity networks for salience processing and executive control. *J. Neurosci.* 27, 2349–2356. doi: 10.1523/JNEUROSCI.5587-06.2007
- Seghier, M. L. (2008). Laterality index in functional MRI: methodological issues. *Magn. Reson. Imaging* 26, 594–601. doi: 10.1016/j.mri.2007.10.010
- Sha, S. J., Takada, L. T., Rankin, K. P., Yokoyama, J. S., Rutherford, N. J., Fong, J. C., et al. (2012). Frontotemporal dementia due to C9ORF72 mutations: clinical and imaging features. *Neurology* 79, 1002–1011. doi: 10.1212/WNL.0b013e318268452e
- Shattuck, D. W., Mirza, M., Adisetiyo, V., Hojatkashani, C., Salamon, G., Narr, K. L., et al. (2008). Construction of a 3D probabilistic atlas of human cortical structures. *Neuroimage* 39, 1064–1080. doi: 10.1016/j.neuroimage.2007.09.031

- Šimundić, A. M. (2009). Measures of diagnostic accuracy: basic definitions. *EFIFCC* 19, 203–211.
- Snowden, J. S., Adams, J., Harris, J., Thompson, J. C., Rollinson, S., Richardson, A., et al. (2015). Distinct clinical and pathological phenotypes in frontotemporal dementia associated with MAPT, PGRN and C9orf72 mutations. *Amyotroph. Lateral Scler. Frontotemporal Degener.* 16, 497–505. doi: 10.3109/21678421.2015.1074700
- Stoodley, C. J., and Schmahmann, J. D. (2010). Evidence for topographic organization in the cerebellum of motor control versus cognitive and affective processing. *Cortex* 46, 831–844. doi: 10.1016/j.cortex.2009.11.008
- Watanabe, Y., and Funahashi, S. (2012). Thalamic mediodorsal nucleus and working memory. *Neurosci. Biobehav. Rev.* 36, 134–142. doi: 10.1016/j.neubiorev.2011.05.003
- Whitwell, J. L., Weigand, S. D., Boeve, B. F., Senjem, M. L., Gunter, J. L., DeJesus-Hernandez, M., et al. (2012). Neuroimaging signatures of frontotemporal dementia genetics: C9ORF72, tau, progranulin and sporadics. *Brain* 135, 794–806. doi: 10.1093/brain/aws001
- Yokoyama, J. S., and Rosen, H. J. (2012). Neuroimaging features of C9ORF72 expansion. *Alzheimers Res. Ther.* 4:45. doi: 10.1186/alzrt148
- Young, K. A., Manaye, K. F., Liang, C., Hicks, P. B., and German, D. C. (2000). Reduced number of mediodorsal and anterior thalamic neurons in schizophrenia. *Biol. Psychiatry* 47, 944–953. doi: 10.1016/S0006-3223(00)00826-X

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.

The handling Editor declared a past co-authorship with one of the authors JL.

Copyright © 2018 Schönecker, Neuhofer, Otto, Ludolph, Kassubek, Landwehrmeyer, Anderl-Straub, Semler, Diehl-Schmid, Prix, Vollmar, Fortea, Deutsches FTL-D-Konsortium, Huppertz, Arzberger, Edbauer, Feddersen, Dieterich, Schroeter, Volk, Fließbach, Schneider, Kornhuber, Maler, Prudlo, Jahn, Boeckh-Behrens, Danek, Klopstock and Levin. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Amyloid Associated Intermittent Network Disruptions in Cognitively Intact Older Subjects: Structural Connectivity Matters

Susanne G. Mueller^{1,2*} and Michael W. Weiner^{1,2}

¹ Center for Imaging of Neurodegenerative Diseases, San Francisco Veterans Affairs Medical Center, San Francisco, CA, United States, ² Department of Radiology and Biomedical Imaging, University of California, San Francisco, San Francisco, CA, United States

OPEN ACCESS

Edited by:

Stefan Teipel,
Deutsche Zentrum für
Neurodegenerative Erkrankungen e. V.
(DZNE), Germany

Reviewed by:

Betty M. Tijms,
VU University Medical Center,
Netherlands
Pew-Thian Yap,
University of North Carolina at Chapel
Hill, United States

*Correspondence:

Susanne G. Mueller
susanne.mueller@ucsf.edu

Received: 30 June 2017

Accepted: 06 December 2017

Published: 19 December 2017

Citation:

Mueller SG and Weiner MW (2017)
Amyloid Associated Intermittent
Network Disruptions in Cognitively
Intact Older Subjects: Structural
Connectivity Matters.
Front. Aging Neurosci. 9:418.
doi: 10.3389/fnagi.2017.00418

Observations in animal models suggest that amyloid can cause network hypersynchrony in the early preclinical phase of Alzheimer's disease (AD). The aim of this study was (a) to obtain evidence of paroxysmal hypersynchrony in cognitively intact subjects (CN) with increased brain amyloid load from task-free fMRI exams using a dynamic analysis approach, (b) to investigate if and how hypersynchrony interferes with memory performance, and (c) to describe its relationship with gray and white matter connectivity. Florbetapir-F18 PET and task-free 3T functional and structural MRI were acquired in 47 CN (age = 70.6 ± 6.6), 17 were amyloid pos (florbetapir SUVR > 1.11). A parcellation scheme encompassing 382 regions of interest was used to extract regional gray matter volumes, FA-weighted fiber tracts and regional BOLD signals. Graph analysis was used to characterize the gray matter atrophy profile and the white matter connectivity of each subject. The fMRI data was processed using a combination of sliding windows, graph and hierarchical cluster analysis. Each activity cluster was characterized by identifying strength dispersion (difference between pos and neg strength) their maximal and minimal pos and neg strength rois and by investigating their distribution and association with memory performance and gray and white matter connectivity using spearman rank correlations (FDR $p < 0.05$). The cluster analysis identified eight different activity clusters. Cluster 8 was characterized by the largest strength dispersion indicating hypersynchrony. Its duration/subject was positively correlated with amyloid load ($r = 0.42$, $p = 0.03$) and negatively with memory performance (CVLT delayed recall $r = -0.39$ $p = 0.04$). The assessment of the regional strength distribution indicated a functional disconnection between mesial temporal structures and the rest of the brain. White matter connectivity was increased in left lateral and mesial temporal lobe and was positively correlated with strength dispersion in the cross-modality analysis suggesting that it enables widespread hypersynchrony. In contrast, precuneus, gray matter connectivity was decreased in the right fusiform gyrus and negatively correlated with high degrees of strength dispersion suggesting that progressing gray matter atrophy could prevent the generation of paroxysmal hypersynchrony in later stages of the disease.

Keywords: amyloid, intermittent, functional connectivity, cognitively intact, hypersynchrony, resting state fMRI, DTI, gray matter map

INTRODUCTION

Abnormal functional connectivity measured by task-free fMRI is one of the earliest manifestations of amyloid positivity in cognitively intact subjects (e.g., Sperling et al., 2009; Sheline et al., 2010; Jack et al., 2013; Wang et al., 2013; Brier et al., 2014; Steininger et al., 2014; Jones et al., 2016). Findings in animal models suggest that the early preclinical phase of Alzheimer's disease (AD) before the onset of cognitive impairment is characterized by hypersynchrony, i.e., an increased tendency for synchronous firing of larger than normal neuron populations that has been linked to increased connectivity or hyperconnectivity in task-free fMRI (Palop and Mucke, 2016; Shah et al., 2016). The findings in human studies in this early preclinical phase are far from consistent with some studies findings increased connectivity (e.g., Lim et al., 2014; Matura et al., 2014; Jiang et al., 2016; Schultz et al., 2017; Sepulcre et al., 2017) and others decreased connectivity (e.g., Sheline et al., 2010; Wang et al., 2013; Steininger et al., 2014; Elman et al., 2016) or both but in different regions (Mormino et al., 2011). Differences between study populations, techniques used to assess functional connectivity, regions investigated etc., undoubtedly contribute to these conflicting findings. However, it is also possible that these discrepancies reflect the true nature of abnormal functional connectivity at this early stage. In all these studies it is usually assumed that hyperconnectivity or hypoconnectivity are sustained. An alternative explanation is that hyper- and/or hypoconnectivity are paroxysmal and therefore that the observation of one or the other just reflects the preferred state of the brain at the time of the exam. There is evidence that this is indeed the case for the early stage hyperconnectivity. For example it has been shown in animal models of AD and also in patients suffering from familial AD that high levels of amyloid and amyloid precursor protein can cause intermittent neuronal hyperactivity in form of epileptic discharges (Palop et al., 2007; Palop and Mucke, 2010; Busche et al., 2012; Grienberger et al., 2012; Mucke and Selkoe, 2012; Talantova et al., 2013; Vossel et al., 2013; Kellner et al., 2014; Born, 2015; Stargardt et al., 2015). Only about 2% of the patients diagnosed with mild cognitive impairment (MCI) or dementia due to sporadic AD show overt epileptiform discharges during routine EEG recordings though. Intermittent unspecific EEG abnormalities, e.g., episodic focal or diffuse slowing of the background activity, indicative for low level focal or diffuse paroxysmal hypersynchrony are more common in MCI and AD and can be found in 20–45% of the routine EEGs (Liedorp et al., 2009, 2010; Kramberger et al., 2013). If paroxysmal hypersynchrony severe enough to be detected by routine EEG occurs in the more advanced stages, it is very well-possible that it is already present in a more subtle form in the preclinical stage and contributes to the conflicting resting state findings depending on its severity and frequency in the study population. Furthermore, complex interictal epileptiform discharges as well as diffuse unspecific EEG abnormalities have been associated with impaired cognitive performance (Smits et al., 2011; Kleen et al., 2013) which raises the possibility that also more subtle types of hypersynchrony in the preclinical stages of AD could already have a negative impact on cognition.

The traditional type of task-free fMRI analysis calculates functional connectivity from the correlation of the BOLD signal fluctuations across the whole acquisition time and thus makes the implicit assumption that these fluctuations stay stable during this time. In diseases that are known or suspected to be associated with paroxysmal events short, infrequent hyperconnectivity phases are likely to be canceled out by longer phases of physiological activity or cause unspecific connectivity disturbances. Studies interested in investigating the dynamic behavior of the BOLD signal therefore use modifications of the traditional stationary approach of which the sliding window approach is one of the most commonly used (Hutchison et al., 2013).

The overall objective of this study was to use a combination of sliding windows with graph and cluster analysis to seek evidence for paroxysmal focal or diffuse hyperconnectivity suggesting paroxysmal low level hypersynchrony in cognitively normal elderly subjects with and without increased brain amyloid load and to characterize its spatial pattern. The second objective was to investigate if these phases could affect cognition by correlating their duration with memory performance and by investigating to what degree lateral and mesial temporal lobe structures, i.e., structures not only known to be involved in memory processes but also to be affected early in AD, participate in these phases. It was assumed that hyperconnectivity phases are more likely to have a negative impact on memory performance if they last longer and if this activity interferes with mesial temporal lobe connectivity. The third and last objective was to investigate a potential relationship between structural (gray and white matter) connectivity and the severity of the functional hyperconnectivity phases. It was assumed that structural connectivity had to be intact to enable network hypersynchrony, i.e., that gray matter atrophy due to synapse and/or neuron loss would prevent the brain from generating the type of abnormal firing associated with hypersynchrony and that white matter damage would prevent the synchronization of remotely abnormally firing regions.

METHODS

Study Population

A total of 47 elderly (60 years and older), cognitively intact subjects who were recruited from local Memory Clinics and the community with flyers and advertisements in local newspapers participated in this study. Exclusion criteria included any poorly controlled medical illness (untreated diabetes, hypertension, thyroid disease) and/or use of medication or recreational drugs that could affect brain function, a history of brain trauma, brain surgery or evidence for ischemic events (stroke but not white matter hyperintensities or small lacunes) and skull defects on the MRI. Normal cognitive functioning was assessed by a battery of standard tests that included the Mini Mental State Examination (MMSE), Clinical Dementia Rating (CDR), California Verbal Learning Test II (CVLT-II), the Wechsler Adult Intelligence Scale III (WAIS III, digit symbol, matrix reasoning), and the Delis-Kaplan Executive Function System (DKEF, trail making, verbal fluency, design fluency). From this battery, three CVLT-II subtests Short Free Discriminability and Delayed

Recall Discriminability, were chosen to assess the impact of amyloid-associated hyperactivity on cognitive function because an association between these measures and brain structure had been demonstrated in a previous study in this cohort (Mueller et al., 2011). Please see **Table 1** for demographic details of the final study population. All participants underwent structural and functional MR imaging and had a florbetapir exam to determine the amyloid beta plaque load. The study was approved by the committees of human research at the University of California, San Francisco (UCSF) and VA Medical Center San Francisco, and written informed consent was obtained from all subjects according to the Declaration of Helsinki.

PET

Florbetapir F18 PET exams were acquired either at the VA Medical Center, San Francisco on a GE Discovery 690 PET scanner or at the China Basin Campus of the University of California, San Francisco on a GE Discovery STE VCT PET system. The participants were injected with 10 mCi (370 MBq) of florbetapir followed by 10 min PET acquisition 50 min later. The images were reconstructed and normalized to a florbetapir template on which several gray matter regions of interest known to be vulnerable to amyloid deposition in the temporal and parietal lobes, precuneus and anterior and posterior cingulate were labeled. The mean count from each of these regions was extracted and regional Standard Uptake Value Ratios (SUVR) were calculated using whole cerebellum as the reference region. The global SUVR was calculated by averaging the SUVRs from all cortical labels. Participants with a global SUVR equal or higher than 1.10 were considered to be amyloid positive.

MR Acquisition

All images were acquired on a Siemens Skyra 3T MR system equipped with a 20 channel receive coil. The following sequences were obtained as part of a larger research protocol. (1) T1-weighted gradient echo MRI (MPRAGE) of entire brain, TR/TE/TI = 2300/2.96/1,000 ms, $1.0 \times 1.0 \times 1.0 \text{ mm}^3$ resolution, acquisition time = 5.30 min for tissue segmentation. (2) PD/T2 weighted 2D turbo spin-echo sequence, TR = 3,210, TE1/2 = 101/11 ms, $1.0 \times 1.0 \times 3.0 \text{ mm}^3$ resolution, acquisition

time: 3.43 min, for co-registration between T1 and EPI data. (3) 2D gradient echo EPI sequence TR/TE = 3,000/30 ms, flip angle = 80, $2.5 \times 2.5 \times 3 \text{ mm}$ resolution, no gaps, acquisition time = 8.00 min for dynamic task-free analysis. Subjects were instructed to refrain from caffeinated beverages on the day of the exam, to close their eyes and to relax but stay awake and think of nothing in particular during the scan. (4) EPI-based diffusion weighted imaging (TR/TE, = 7,200/73 ms, $2 \times 2 \times 2$ resolution, 64 diffusion encoding directions with $b = 1,000 \text{ s/mm}^2$, acquisition time: 7.2 min.

Image Processing

Task Free Functional Imaging Data

The first six time frames were discarded to allow the MRI signal to achieve T1 equilibrium. The remaining 154 timeframes/subject underwent slice time correction, motion correction and realignment onto a mean EPI image in the T1 space, spatial normalization using the transformation matrices generated during the warping of the gray matter maps onto the gray matter template with re-sampling to a $1.5 \times 1.5 \times 1.5 \text{ mm}$ resolution. Framewise displacement (Power et al., 2012) was used to assess the motion during the exam. Conn 17a (www.nitrc.org/projects/conn, Whitfield-Gabrieli and Nieto-Castagnon, 2012) a SPM based toolbox for task and task-free fMRI analysis was used for further processing including linear detrending and band pass filtering (0.008–0.09 Hz) with simultaneous denoising. The latter included the aCompCorr routine to reduce the effects of physiological noise (eroded white and csf maps, five components each) and motion regression (six affine motion parameters and six first order temporal derivatives). In addition to that, ART as implemented in the conn preprocessing was used to identify timeframes with motion exceeding a movement threshold of 0.9 mm which ensures that conn disregards these timeframes during the denoising procedure but leaves the original time series intact. No global signal removal was performed since this is known to falsely increase anti-correlations between time series (Murphy et al., 2009). The AICHA atlas was used to extract the denoised mean time series and to estimate the functional connectivity.

Stationary or Time Average Analysis

Timeframes identified as having excessive motion by ART were removed and the functional connectivity matrix calculated. The routines provided by the Brain Connectivity Toolbox (<https://sites.google.com/site/bctnet>), in particular the weight conserving measures “strength,” was used for this purpose (Rubinov and Sporns, 2011). Weight conserving measures have the advantage that they can be applied to fully connected networks, i.e., it is not necessary to define an arbitrary threshold to generate the type of sparse network required by the more commonly used non-weighted equivalents degree. Strength is defined as the sum of weights of links connected to a node or roi. A roi has a high positive strength if its BOLD fluctuations are positively correlated with a large number of that of other rois and a high negative strength if its BOLD fluctuations are negatively correlated with a large number of that of other rois. Positive (Spos) and negative (Sneg) nodal strength and nodal strength dispersion (Sdisp), defined as

TABLE 1 | Subject characteristics.

| | Amyloid negative <i>n</i> = 30 | Amyloid positive <i>n</i> = 17 |
|--|-----------------------------------|-----------------------------------|
| Age | 70.2 (6.3) | 72.1 (6.3) |
| SUVR | 1.02 (0.06) | 1.25 (0.11)* |
| ApoE4 pos/neg | 7/23 | 8/11 |
| CDR | 0.0 (0.00) | 0.00 (0.0) |
| MMSE | 29.8 (0.5) | 29.6 (0.8) |
| CVLT-II immediate recall discriminability | 2.3 (0.4) | 2.4 (0.7) |
| CVLT-II short free recall discriminability | 2.6 (0.7) | 2.5 (0.9) |
| CVLT-II delayed recall discriminability | 2.6 (0.7) | 2.5 (0.9) |
| DKEF verbal fluency | 13.9 (3.33) | 13.9 (3.4) |
| Digit symbol | 65.6 (17.1) | 64.1 (12.2) |

difference between nodal positive and negative strength, were used to assess the effects of group, age, and SUVR on functional connectivity. Sdisp increases when Spos increases and Sneg simultaneously decreases, i.e., shows a strength profile consistent with hyperconnectivity or hypersynchrony, and was therefore used as a proxy of hypersynchrony.

Dynamic Analysis

A sliding windows approach was used to explore temporal variations of functional connectivity. Based on observations that robust estimations of the functional connectivity without loss of potentially interesting fluctuations are possible with window sizes around 30–60 s (Hutchison et al., 2013), a window with the size of 45 s (15 timeframes) that was advanced with increments of one TR along the artifact corrected time series was chosen resulting in 138 windows/subject or 6,486 windows for all 47 subjects that were converted into 6,486 correlation matrices using Pearson correlation (cf **Figures 1A–C**).

Graph analysis was again used to describe the interactions between the different nodes in each window (please see **Figure 1D**). The strength outputs for each window were combined to obtain a map showing the fluctuations of pos and

neg strength over the whole acquisition time for each roi for each subject (**Figure 1E**) and then concatenated across subjects (**Figure 1F**) to obtain population maps of pos and neg strength (**Figure 1F**). The nodal positive and negative strength in each window of this population map were converted into z-scores using mean and standard deviation of the nodal strength of the amyloid neg CI as reference with the following formula: strength z-score of node x in window $n = \text{strength of node } x \text{ in window } n - \text{mean of strength of node } x \text{ from all windows in reference data set} / \text{standard deviation of strength of node } x \text{ from all windows in reference data set}$. The thus calculated nodal z-scores/window were averaged over all nodes to obtain global positive and negative strength z-scores for each window in each subject (**Figure 1G**). Hierarchical cluster analysis (Ward's minimum variance methods with the cubic clustering criterion to identify optimal cluster number) was used to identify different global negative and positive strength profiles in amyloid pos and neg subjects (optimal cluster number = 11, please see **Figure 1H**). The output generated by ART was used to identify windows with motion outliers in the real data and to calculate the percentage of motion outliers for each cluster. Clusters 9–11 consisted of more than 50% of motion outliers (range 75–100%) and therefore

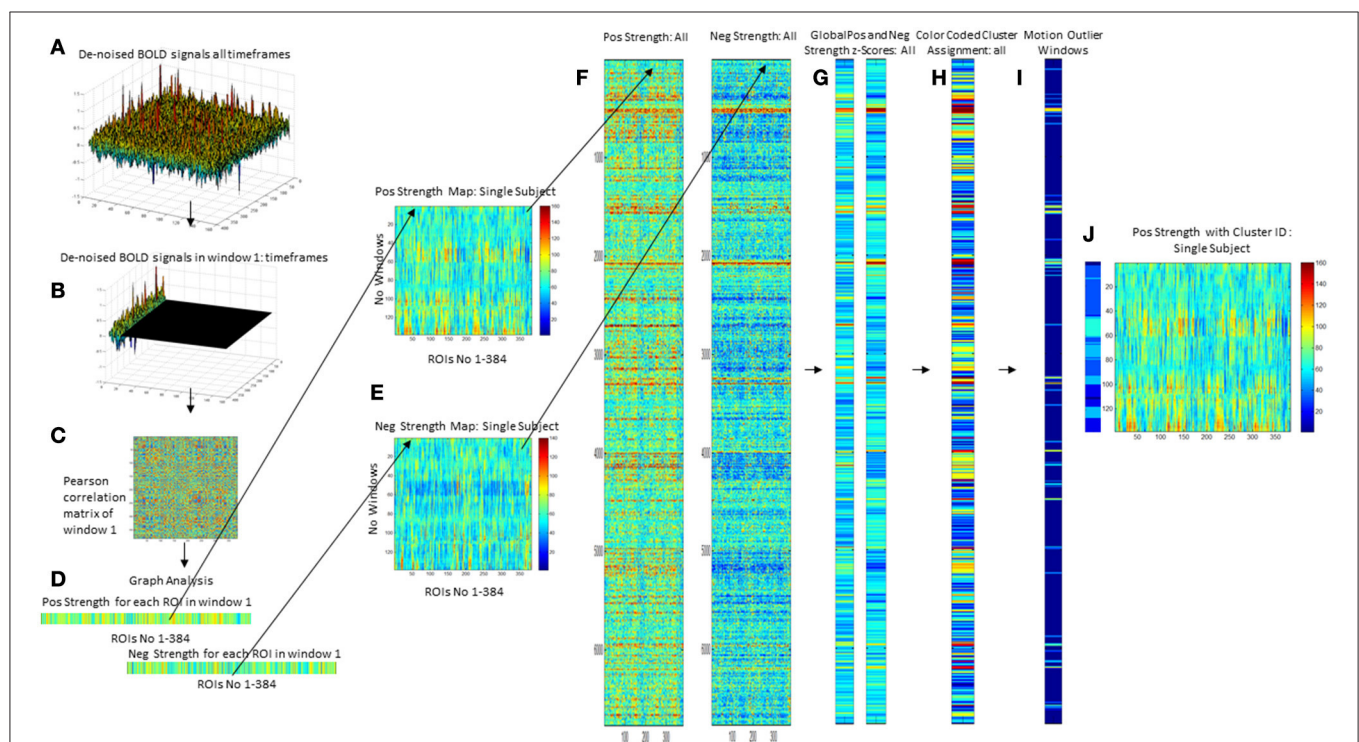


FIGURE 1 | Overview of processing steps. An example of denoised BOLD signal timeseries of an amyloid neg subject is depicted on the upper right side (**A**). The timeseries is divided into overlapping 45 s long epochs or windows using a sliding windows approach (**B**), and a correlation matrix calculated using Pearson correlation (**C**). Graph analysis is used to describe the interactions between the different ROIs for each window (**D**) that are combined to obtain maps that depict the fluctuations of positive (pos) and negative (neg) strength over the subject's whole timeseries (**E**). For the analysis the strength maps of all 47 subjects were combined (**F**), converted into z-scores using the mean and std from the amyloid neg CN group as a reference from which the mean from all nodes is calculated for pos strength and neg strength (**G**). This is used as the input for a hierarchical cluster analysis to identify windows with a similar global pos and neg strength profile (**H**). The output from ART is used to identify windows with excessive motion or global signal fluctuations (**I**) and to eliminate them from all further analyses. The right-most panel shows the cluster assignment for the sample subject (**J**). Each of the 138 windows in this subjects strength maps has been assigned to a cluster which allows to determine how many different clusters or activity states occur in each subject and how long they last (duration = no of windows assigned to the cluster/activity state).

were considered to represent “motion clusters” and were not further evaluated. This also eliminated windows that despite not meeting the ART threshold for motion outliers themselves had a similar graph analytical profile as windows that did meet that threshold and were therefore likely to be affected by subthreshold motion. Forty-one percent of the windows assigned to cluster 1 had been identified as motion outliers and therefore cluster 1 was considered as “motion contaminated.” All other clusters had 20% or less motion outliers and were together with cluster 1 fully evaluated after excluding all motion outlier windows. Eliminating windows with excessive motion results in a more rigorous elimination of motion artifacts than just eliminating the motion affected timeframe alone because it also eliminates timeframes with subthreshold motion that usually accompany timeframes with suprathreshold motion (**Figures 1I,J**).

The last step was to investigate if certain clusters tended to occur together. This was done by calculating the “cluster neighborhood” or the frequency by which a window that had been assigned to cluster A were next to a window assigned to one of the other clusters in a data set. The frequencies with which the other clusters appeared were compared with Fisher’s exact test ($p < 0.05$ with Bonferroni correction for multiple comparisons) to identify those cluster(s) that were significantly more often found in the neighborhood of cluster A than others clusters. **Table 2** summarizes the global strength profiles of the eight clusters expressed as z-scores. Clusters 1–8 were further characterized by investigating the following features.

1. At the global level by investigating the relationship between SUVR and memory performance (short and delayed recall of the CVLT) and the time (counts of windows assigned to cluster/subject) each strength profile or activity type could be observed in an individual CN.
2. At the nodal level by determining mean nodal strength dispersion for each cluster to investigate differences between the distribution of brain regions with high and low strength dispersion between clusters. Given the a priori hypotheses re memory, the distribution within the medial temporal regions (hippocampus, parahippocampus and fusiform gyrus) and lateral temporal regions (superior, middle and inferior temporal gyrus) was of particular interest. Finally, the relationship between the time a particular strength dispersion

profile was observed in an individual CN and SUVR and memory performance was investigated.

Volumetric Imaging Data

The T1 images were segmented using the new segmentation algorithm as implemented in SPM12. The gray matter maps were warped onto a symmetrical gray matter atlas in MNI space while preserving the total amount (modulation) using SPM12’s DARTEL routine and corrected for intracranial volumes (ICV) using the individual’s combined gray/white/csf volumes. The resulting gray matter maps were then converted into z-score maps using the mean and standard deviation of the gray matter maps of 32 healthy young controls (mean age: 28.2 ± 6.7) who had been studied with the same sequence on the same magnet and whose images had undergone the same processing. The atlas of intrinsic connectivity of homotopic areas (AICHA, Joliot et al., 2015) consisting of 384 homotopic cortical, and subcortical gray matter regions of interest (gm roi) was used to extract the mean z-score values for each gm roi. In order to investigate how the mean gray matter z-score in one region is related to that of other regions the so-called profile similarity index (PSI) was calculated. The PSI between gm roi x and gm roi y was defined as follows:

$$\text{rawPSI} = (\text{cROIA} - \text{mean}_{\text{roi}}) / (\text{abs}((\text{gm roi x} - \text{mean}_{\text{roi}}) - (\text{gm roi y} - \text{mean}_{\text{roi}})))$$

cROIA is either gm roi x or gm roi y whichever is larger, mean_{roi} is mean over all 384 gm rois.

The rawPSI is calculated for each and every combination between two gm rois resulting in a 384×384 matrix for every subject. RawPSI-values exceeding the 95 percentile of all PSI-values in the map are replaced by the PSI-value at the 95 percentile to remove outliers caused by a difference of 0 or very small differences between two gm rois. The rawPSI map is then converted into the final PSI map by multiplying it with a normalization term n defined as $n = 1/(\text{range of all raw PSI in map})$. A negative PSI indicates a gray matter loss and a positive PSI a gray matter increase relative to the subject’s mean z-score. The resulting PSI map of a healthy older individual is determined by atrophic changes due to normal aging. Although an individual PSI map is defined by individual anatomical features, it is assumed to share many features with PSI maps of other healthy

TABLE 2 | Global strength profiles of all non-motion clusters.

| Cluster ID | NoWindows | meanSposZ | meanSnegZ | Median fwd | Neighbors | No amyloid neg vs. pos | Strength profile type |
|------------|-----------|-----------|-----------|------------|----------------|------------------------|-----------------------|
| 1 | 163 | −0.116 | 0.438 | 0.258 | 2, 3 | 19/9 | Non-balanced negative |
| 2 | 909 | −0.097 | 0.086 | 0.241 | 4 | 28/17 | Non-balanced negative |
| 3 | 926 | −0.490 | 0.155 | 0.192 | 4, 2 | 23/16 | Non-balanced negative |
| 4 | 1,202 | −0.332 | −0.069 | 0.224 | 5, 2, 3 | 30/16 | Balanced |
| 5 | 928 | −0.182 | −0.259 | 0.232 | 4, 7 | 28/17 | Balanced |
| 6 | 452 | 0.362 | −0.197 | 0.236 | 7, 2 | 23/12 | Non-balanced negative |
| 7 | 775 | 0.102 | −0.476 | 0.236 | 5, 8, 6 | 22/14 | Non-balanced negative |
| 8 | 302 | 0.618 | −0.727 | 0.253 | 7, 6 | 14/7 | Non-balanced negative |

fwd, framewise displacement. please see text for details. Neighboring clusters, bold, cluster that appears most often together with this cluster, non-bold clusters also appearing together with this cluster but less often than bold cluster.

elderly subjects. A pathological process though, e.g., gray matter loss in medial temporal structures due to tau pathology, will change the appearance of the PSI map. Graph theory is used to characterize each subject's PSI map. A gm roi has a high strength if it has experienced a similar degree of gray matter loss as the majority of the other gm rois and a low strength if there are only few other gm rois with similar gray matter atrophy. This project focused on differences of negative strength (sNeg) since the focus of this study was on additional gray matter atrophy due to a pathological process. The influence of group (amyloid pos CN vs. amyloid neg CN), age and SUVR on nodal negative gray matter strength was investigated.

Diffusion Weighted Imaging Data

ExploreDTI (Leemans et al., 2009) was used to process the DTI data. After correction for motion and eddy-current induced geometric distortions (Leemans and Jones, 2009; Irfanoglu et al., 2012), diffusion tensors were calculated using a non-linear regression procedure. Whole brain fiber tracking was done using a deterministic streamline method and fiber pathways reconstructed by defining seed points uniformly throughout the brain (FA thresholds = 0.2, angle threshold = 30 degrees, step size 1). The reconstructed fiber tracts were parcellated using the AICHA parcellation that had been warped onto each subject's B0 map in subject space. White matter connectivity maps were generated by identifying the number fiber connections passing through both rois and extracting the mean FA from those fibers. Graph analysis using weight conserving measures, i.e., positive strength, was used to describe each subject's white matter connectivity map at the nodal level (wm rois) and to assess group differences (amyloid pos CN vs. amyloid neg CN) and the influence of age and SUVR on white matter connectivity. White matter connectivity is denoted with cFA.

Cross-Modality Analysis

The influence of white and gray matter connectivity on stationary and dynamic functional strength, dispersion was assessed by correlating each nodes strength dispersion with that of its own white (cFA positively correlated with Sdisp) and gray matter (sNeg negatively correlated with Sdisp) connectivity and with the white and gray matter connectivity of every other node. To reflect the temporal aspect of the dynamic analysis, nodal Sdisp was weighted by the time during which this cluster was observed in the individual before correlating it with white/gray matter connectivity measures. To facilitate the interpretation of these cross-modality analyses, the brain was divided into 20 regions (left and right, lateral frontal, medial frontal, cingulate, insula, lateral temporal, medial temporal, lateral parietal, medial parietal, occipital, and subcortical) and significant correlations between rois within a region or with rois in other regions counted (regional cross-modality connectivity matrix). To account for the different number of rois within a region, the connectivity within or between regions was expressed as a ratio (count of significant correlations/no of rois in region with fewer rois). The overall connectivity of each region was determined by summing up the region's entries in the cross-modality connectivity matrix along the x and y axis (region connectivity summary). Upper and

lower 99% confidence intervals were calculated for each region summary to identify regions with an increased (>99% upper confidence interval) and decreased (<99% lower confidence interval) cross-modality interaction by bootstrapping (10,000 iterations).

Statistics

Stationary fMRI analysis/white matter and gray matter connectivity analyses: two-tailed Spearman correlation analyses corrected (FDR $p < 0.05$) for multiple comparisons were used to assess each nodes strength with age and SUVR.

- Cross-modality analyses: one-tailed Spearman correlation analyses corrected for multiple comparisons (FDR $p < 0.05$) were used to identify significant correlations between rois.
- Dynamic analysis: Mann Whitney and one tailed Spearman correlation tests were used to test for differences of the occurrences of different clusters between amyloid pos and amyloid neg CN and to investigate the relationship between cluster counts and SUVR, age and cognitive function. JMP 12.1.0 was used for the cluster analysis, the statistical toolbox in Matlab was used for all other statistical analyses.

RESULTS

Functional Connectivity

Stationary Analysis

Figure 2A displays rois whose nodal strength correlated with age and SUVR. Age was significantly negatively correlated with Spos of rois in the left supramarginal gyrus and right cuneus and positively with a roi in the right putamen. Sneg was positively correlated with age in rois in the left superior frontal, medial temporal and right thalamus and in the left and right anterior insula and mesial prefrontal region. Sdisp was negatively correlated with age in rois of the left supramarginal gyrus and middle temporal gyrus, right mesial prefrontal gyrus and cingulum and bilateral anterior insula.

Spos and Sdisp were positively correlated with SUVR in the left anterior superior temporal gyrus and Sneg was negatively correlated with SUVR in the left orbital-frontal, anterior insula, parieto-occipital and precuneus region and the right superior temporal region. There were no significant correlations between stationary Spos, Sneg, or Sdisp and short or delayed recall discriminability. Taken together, increasing age induced a shift from positive to negative strength, while SUVR had the opposite effect and demonstrated evidence for a SUVR associated focal hypersynchrony in the left superior temporal gyrus.

Dynamic Analysis

The number of different clusters or types of activity observed in a subject ranged between 4 and 8 (median: 6). Please see also **Table 2**. Clusters 1–5 were characterized by a relative decrease of Spos that was accompanied by a decrease of Sneg in clusters 4 and 5 and an increase of Sneg in clusters 1–3. Clusters 6–8 were characterized by an increase of Spos and a decrease of Sneg. Clusters 1–3 and clusters 6–8 represent unbalanced states in which one connectivity type dominates. In the case of clusters 1–3 the balance is shifted toward negative strength indicating the

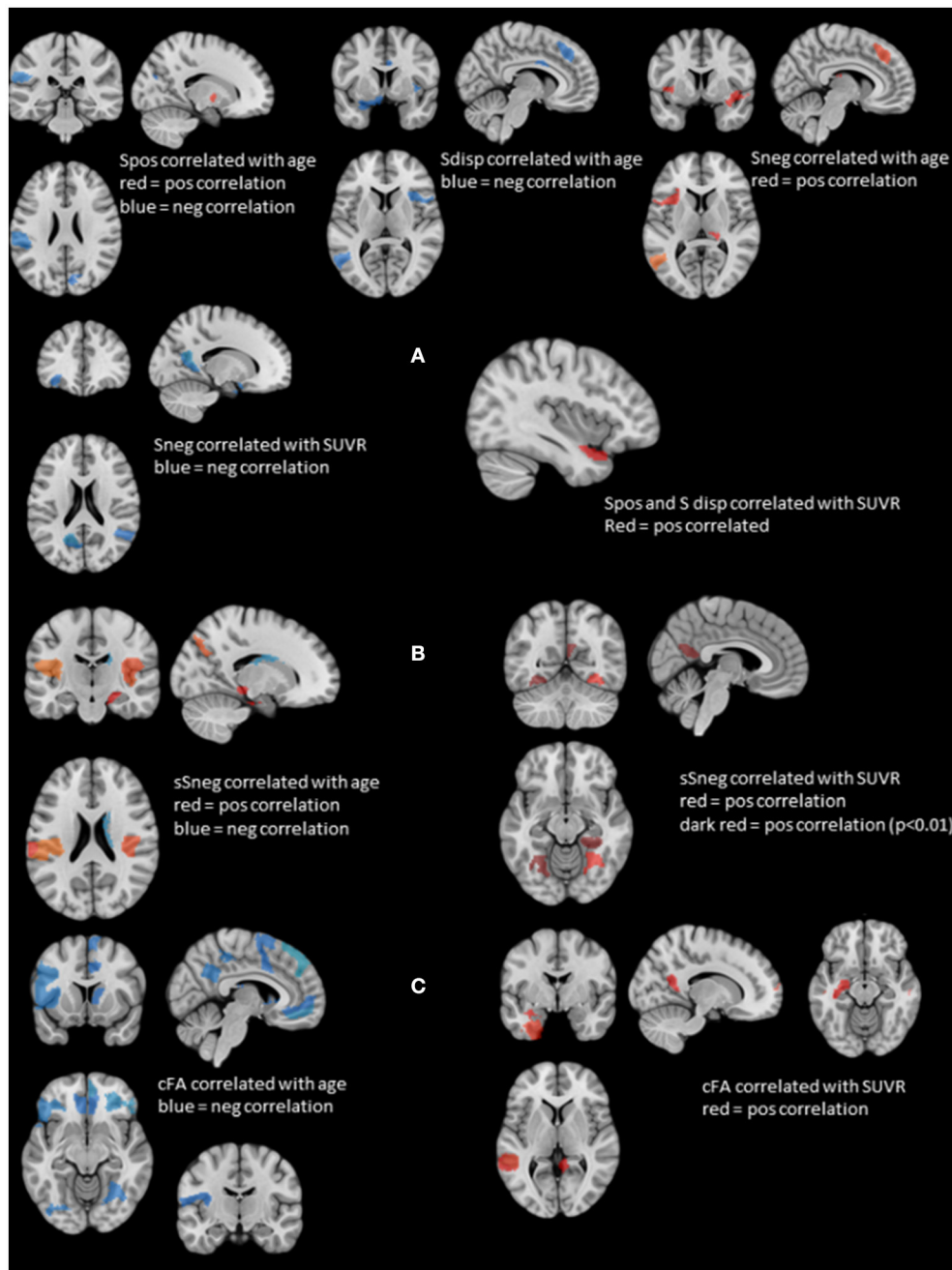


FIGURE 2 | Summary of the structural and stationary connectivity analyses. Regions with positive correlations are displayed in red, those with negative correlations in blue. **(A)** Summarizes the findings of the stationary functional analysis for each of the measures, i.e., positive strength (Spos), negative strength (Sneg) and strength dispersion (Sdisp, calculated as difference between Spos and Sneg) with age and SUVR. **(B)** Displays the findings for atrophy-based gray matter connectivity (please see Methods in text body for details). On the left side regions whose negative structural strength (sSneg) is positively correlated with age, and on the right side regions whose sSneg is positively correlated with SUVR. There was no overlap between regions affected by age and those affected by amyloid load. **(C)** Displays the findings for white matter connectivity (please see Methods in text body for details). cFA is FA weighted stream line count connecting two regions. Correlations with age are displayed on the left and those with SUVR on the right. Again, there was no overlap between regions correlated with age and those correlated with SUVR.

presence of a large number of negative correlations between rois. In clusters 6–8 the balance is shifted toward positive strength consistent with predominately positive correlations between rois. Clusters 4 and 5 represent balanced states during which Spos

and Sneg are both slightly reduced. Clusters with similar strength profiles tended to be neighbors indicating that those with less pronounced strength profile represent a transition from one strength profile into another, e.g., from unbalanced negative to

balanced. Cluster 8 had the highest Spos and lowest Sneg and consequently the largest Sdisp of all clusters, i.e., represented a highly unbalanced state favoring positive correlations which is consistent with a strength profile of hypersynchrony. Cluster 8 was also the only cluster whose window counts/subject were positively correlated with a subject's SUVR ($r = 0.42$, $p < 0.03$) and negatively with a subject's memory performance (CVLT short free recall, $r = -0.43$, $p < 0.03$, delayed recall $r = -0.39$, $p < 0.04$). Please see **Supplementary Table 1** for other clusters.

Figure 3 displays the mean nodal Spos (yellow-red) and nodal Sneg (green-blue) and maximal (75–100 percentile, red) and minimal (0–25 percentile, blue) Sdisp nodes for each cluster. Although the global Spos and Sneg of most clusters differed (please see **Table 3**), the distribution of the respective maxima and minima was quite similar. Although the extent of involvement varied, the mesial temporal region including the inferior temporal and fusiform gyrus and the orbito-frontal region were identified as minimal Sdisp zones in all eight clusters. The minimal Sdisp zones were caused by a lower Spos but higher Sneg compared to other regions. Additional less consistent minimal Sdisp zones were found the dorso-lateral and mesial frontal cortex. The middle and superior temporal gyrus, lateral temporo-parietal region, lateral and medial occipital gyri without occipital poles, and the precuneus were identified as maximal Sdisp zones in all eight clusters. The mesial superior frontal region, anterior cingulate, the paracentral lobules, and pre- and post-central gyri were identified as additional less consistent maximal Sdisp zones. In accordance with its profile in the cluster analysis, cluster 8 was characterized by the highest global mean Spos and lowest global mean Sneg. This shift from Sneg to Spos at the global level was also observed in the lateral temporal lobe where over 50% of its area were identified as maximal Sdisp zones and only 23% as a medium (25–50 percentile) Sdisp zones which was clearly smaller than in other clusters (please see **Table 3**). The mesial temporal regions showed the opposite pattern, i.e., only 9% of its area were identified as maximal Sdisp regions but over 55% as medium Sdisp zone. This indicates that large parts of the mesial temporal lobe were not able to engage in the high Sdisp activity observed in the lateral temporal lobes and other extratemporal brain regions. This pattern was only found in Cluster 8.

The next step was to investigate a potential relationship between memory performance and the observed type of strength re-distribution. To this purpose windows characterized by larger than normal (above upper 99% confidence interval) maximal Sdisp zones and lower than normal (below lower 99% confidence interval) medium Sdisp zones in the lateral temporal regions and windows characterized by smaller than normal (below lower 99% confidence interval) maximal Sdisp and larger than normal (above upper 99% confidence interval) medium Sdisp zones were identified in each cluster and counted for each subject. In cluster 8, the counts of windows with smaller than usual medium Sdisp zones (25–50 percentile) in the lateral temporal lobe were negatively correlated with short free recall ($r = -0.56$, $p = 0.005$) and delayed recall ($r = -0.53$, $p = 0.008$). The counts of windows with larger than usual maximal Sdisp zones in the lateral temporal lobe were also significantly negatively correlated

with short recall and delayed recall but these correlations did not survive correction for multiple comparisons. The counts with larger than usual medium (25–50 percentile) Sdisp zones in the mesial temporal region were negatively correlated with delayed recall ($r = -0.53$, $p = 0.008$). The correlations between counts of windows with larger than usual medium Sdisp zones for short recall and those for smaller than usual high Sdisp zones were also significantly negatively correlated with short recall and delayed recall but did not survive correction for multiple comparisons. Taken together, the findings indicate that a shift of the strength distribution toward higher Spos with simultaneous decrease of Sneg in the lateral temporal lobe that excludes a large part of the medial temporal lobe could have an adverse effect on memory performance if they occur frequently or over a longer time. None of the other clusters showed this kind of relationship with memory performance.

Structural Connectivity

Figure 2B shows gm rois whose sSneg was significantly correlated with age. Positive correlations indicating increased connectivity with other gray matter regions with similar degrees of age-related atrophy were found in gm rois in the left supramarginal gyrus, right superior temporal, parahippocampal and parieto-occipital regions and bilateral posterior insula. A roi in the right caudate was negatively correlated with age. A single roi in the right fusiform gyrus was positively correlated with SUVR. When a more liberal threshold was used ($p < 0.01$) additional rois with positive correlations between sSneg and SUVR were found in the left and right fusiform gyrus and the left precuneus. Taken together, age and SUVR were both positively correlated with sSneg. Age-related sSneg increases were diffuse while SUVR-related Sneg increases were restricted to fusiform gyrus and precuneus. There was no overlap between rois showing age-related sSneg increases and rois showing SUVR related sSneg increases.

Wm rois whose cFA was negatively correlated with age were found in left and right orbito-frontal, inferior frontal, occipital, lateral parietal, precuneus and cingulate regions, in left precentral, rolandic, middle frontal, anterior insula and right superior and medial frontal, superior temporal, fusiform and thalamus regions. Wm rois with cFA that was positively correlated with SUVR were found in left superior temporal, hippocampus, fusiform regions and right superior frontal and middle temporal regions and bilateral precuneus (cf. **Figure 2C**). Taken together, age was negatively correlated with cFA and age-related cFA decreases were widespread but were more common prefrontal. In contrast, cFA was positively correlated with SUVR. Significant correlations were restricted to rois within the temporal lobes and precuneus. There was no overlap between rois correlated with age and those correlated with SUVR.

Cross-Modal Correlations

Figure 4 shows the regional cross-modality connectivity matrices and the graphical representation of the region connectivity summaries for the stationary analyses and each cluster. All clusters except No. 1 had several regions with above threshold positive correlations between Sdisp and cFA indicating that white

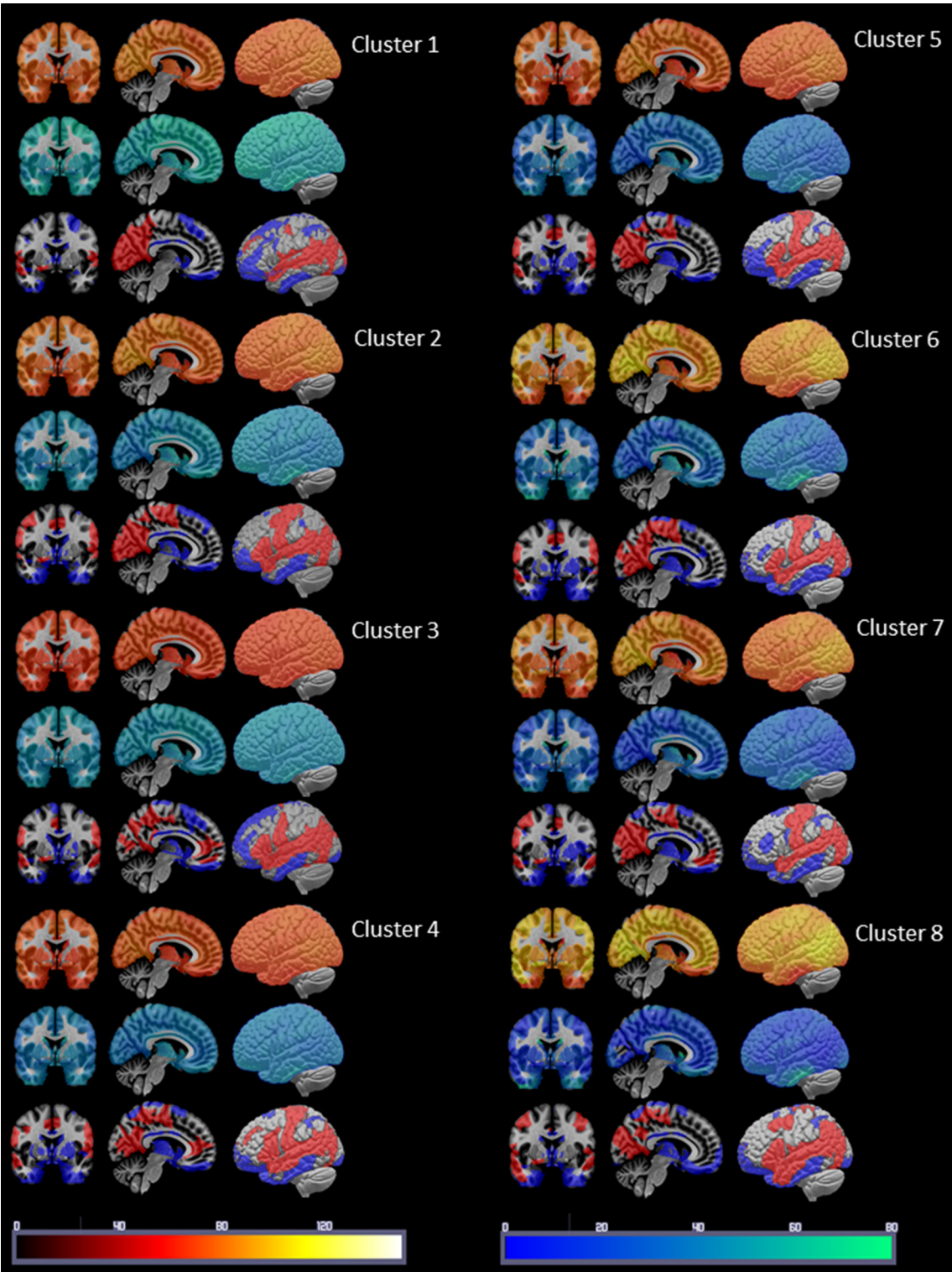


FIGURE 3 | Summary of the cluster characterization. The upper row shows the distribution of Spots in warm colors (please see color bar at the bottom of the figure), the middle row shows the Sneg distribution in cold colors (please see color bar at the bottom) and the lower row the maxima (>75 percentile) of S disp in red and the minima (<25 percentile) in blue. Please see Results section for a description of the findings.

TABLE 3 | Summary of regional strength distribution.

| | Measure | Cluster 1 | Cluster 2 | Cluster 3 | Cluster 4 | Cluster 5 | Cluster 6 | Cluster 7 | Cluster 8 |
|------------------|----------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
| Global | Spos | 76.6 (3.8) | 76.7 (1.9) | 66.9 (1.5) | 70.8 (1.4) | 74.3 (1.0) | 87.7 (3.1) | 81.1 (2.1) | 92.9 (3.5) |
| | Sneg | 64.8 (1.4) | 57.7 (1.5) | 59.1 (1.4) | 54.8 (0.8) | 51.0 (0.6) | 52.1 (2.3) | 46.9 (1.5) | 42.2 (1.8) |
| | % Low S disp | 19.6 | 20.2 | 20.1 | 20.3 | 20 | 21.2 | 21 | 18.9 |
| | % Medium Sdisp | 56.3 | 50.5 | 52.1 | 48.2 | 51.5 | 51.2 | 51.4 | 51.2 |
| | % High S disp | 24.1 | 29.3 | 27.8 | 31.5 | 28.5 | 27.6 | 27.6 | 29.9 |
| Temporal lateral | Spos | 76.5 (5.7)* | 77.0 (5.4)* | 67.1 (3.2)* | 71.2 (4.1)* | 75.0 (6.1)* | 86.5 (9.2)* | 81.0 (4.8)* | 94.1 (8.5) |
| | Sneg | 62.6 (5.0)* | 56.8 (5.1)* | 58.0 (3.1)* | 53.6 (2.8)* | 50.0 (4.4)* | 53.3 (8.3)* | 46.4 (3.5)* | 41.6 (4.4)* |
| | % Low S disp | 11.7 | 22.9 | 19 | 20.5 | 21.9 | 31.4 | 27.2 | 26.5 |
| | % Medium Sdisp | 55.6 | 43.5 | 42.2 | 35.9 | 43.4 | 44.4 | 41.5 | 23.5 |
| | % High S disp | 32.7 | 33.6 | 38.8 | 43.6 | 34.7 | 24.2 | 31.3 | 50.0 |
| Temporal medial | Spos | 75.1 (6.6)* | 74.4 (5.3)* | 65.2 (2.5)* | 68.8 (4.2)* | 71.7 (5.5)* | 84.2 (11.8) | 78.4 (7.1)* | 86.3 (11.2) |
| | Sneg | 63.5 (5.4)* | 57.8 (4.8)* | 58.2 (2.6)* | 54.4 (3.5)* | 51.4 (4.7)* | 54.4 (9.0)* | 48.2 (5.9) | 46.2 (7.6) |
| | % Low S disp | 32.3 | 41.7 | 33.2 | 37.7 | 42.3 | 42.6 | 37.6 | 35.3 |
| | % Medium Sdisp | 46.5 | 37.1 | 48.9 | 49.3 | 44.5 | 38.8 | 43.8 | 55.5 |
| | % High S disp | 21.2 | 21.2 | 17.9 | 13 | 13.2 | 18.6 | 18.6 | 9.2 |

*Significant different compared to cluster 8, global Spos, all clusters except 1 and 2 different, global Sneg all clusters different, pos, pos strength, Sneg, neg strength, S disp, strength dispersion, low, within 0–25 percentile; high, with 75–100 percentile, % coverage of total lat TL or med TL area. Bold highlights S disp behavior unique to cluster 8.

matter connectivity had a role in maintaining Sdisp regardless of the magnitude of Sdisp. The most prominent Sdisp cFA correlations were found for Cluster 3 and 6 with 13, respectively, 14 above threshold regions. Cluster 8 had only three above Sdisp/cFA threshold regions, one of them the left mesio-temporal region. Cluster 8 and 7, i.e., both clusters with an unbalanced positive strength profile and high Sdisp were the only clusters that had several regions with above threshold negative correlations between Sdisp and sSneg indicating an adverse effect of gray matter atrophy on maintaining activity characterized by high S disp.

DISCUSSION

There were two major findings: (1) The dynamic functional connectivity analysis revealed paroxysmal phases of unbalanced activity characterized by a widespread increased strength dispersion, i.e., high pos strength associated with low neg strength, consistent with hypersynchrony in cluster 8. The duration of these phases was positively correlated with amyloid load indicating a relationship between amyloid load and the occurrence and severity of these paroxysmal phases. The widespread shift toward positive strength at the expense of negative strength also included large parts of lateral temporal lobes but mostly spared mesial temporal lobe structures, indicating a state of functional disconnection of the mesial temporal region. The duration of this mesial temporal disconnection state was negatively correlated with memory performance. This paroxysmal widespread hypersynchrony seen in the dynamic analysis was only weakly reflected in the traditional time-averaged analysis that showed decreased neg strength in isolated rois in the frontal, parietal and temporal

lobes and increased pos strength in the superior temporal lobe. There was no association with memory performance in the time averaged analysis. (2) Amyloid load was associated with an increased white matter connectivity in the left lateral and mesial temporal lobe and precuneus and with an increased atrophy related connectivity in the right fusiform gyrus. The findings of the cross-modality analysis suggest that the increased white matter connectivity enabled the brain to maintain the hypersynchrony and that the altered gray matter connectivity in the mesial temporal lobe contributed to the functional disconnection of this region. Taken together, the findings support the notion that increased amyloid load in CN is associated with phases of widespread paroxysmal hyperconnectivity consistent with hypersynchrony and that these phases could have a negative impact on memory. The findings of the structural analysis suggest that this widespread paroxysmal hypersynchrony depends on an intact structural connectivity, i.e., that they either become less widespread or vanish completely when the neurodegenerative process progresses. The following paragraphs will discuss these findings in more detail and attempt to put them into the context of the current knowledge.

The first major finding of this study was that an increasing amyloid load was associated with increasingly longer paroxysmal states characterized by increased positive strength and simultaneously decreased negative strength resulting in a widespread increased strength dispersion. This hyperconnectivity pattern is consistent with a widespread low level hypersynchrony that has also been described in AD animal models at a very early stage of the disease (Busche and Konnerth, 2016; Shah et al., 2016). Although widespread, the strength dispersion had regional maxima in the medium and superior temporal lobes, lateral temporo-parieto-occipital region and precuneus. Regional strength dispersion minima caused

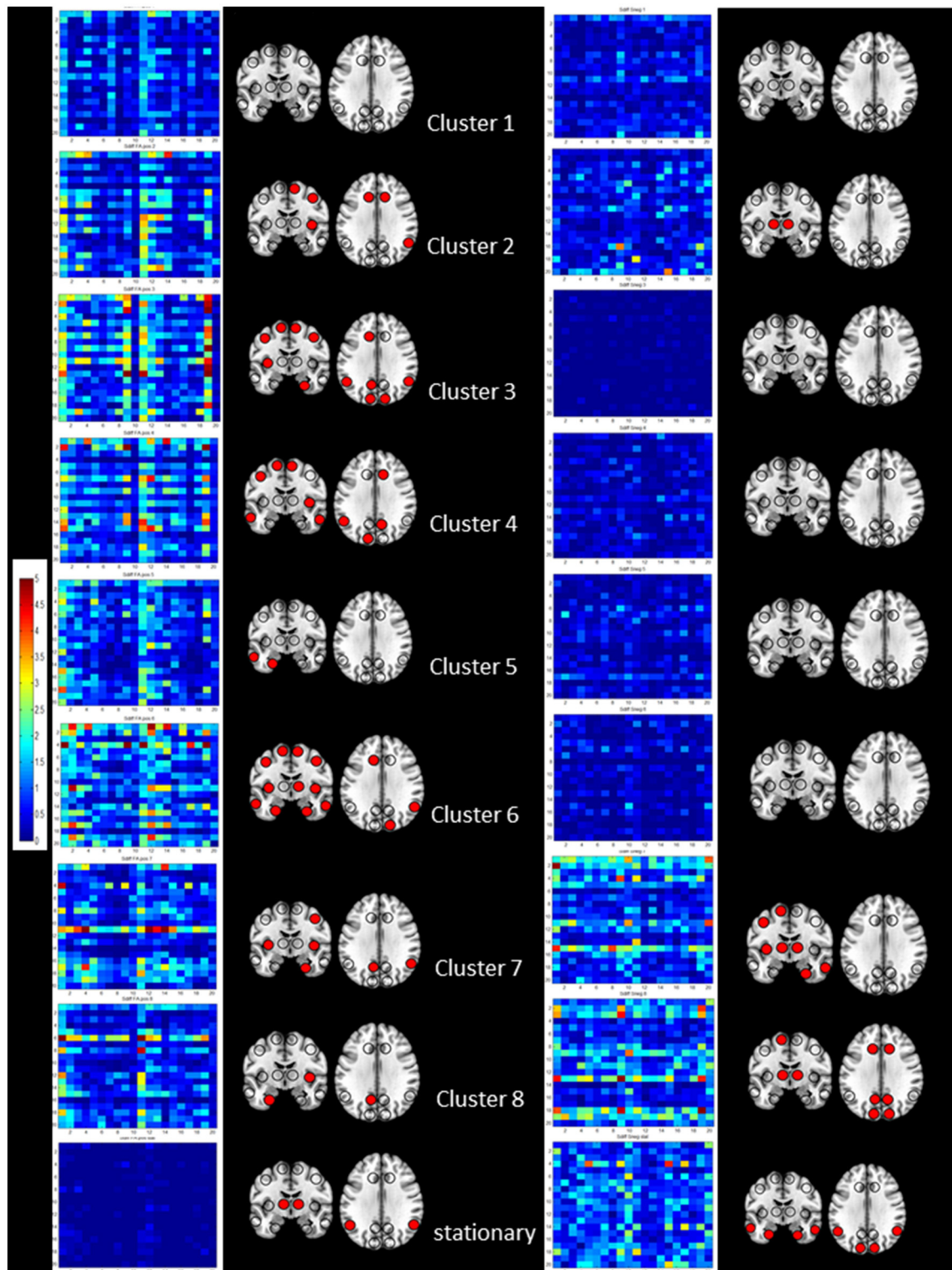


FIGURE 4 | Summary of the results of the cross-modality correlation for each cluster (1–8 from top to bottom) and the stationary analysis (bottom). On the left side significant positive correlations between S disp and cFA are summarized by their regional cross-modality connectivity matrices displaying 20 regions (x from left to right and y top to bottom 1, left lateral frontal, 2, left medial frontal, 3, left cingulate, 4, left insula, 5, left lateral temporal, 6, left medial temporal, 7, left lateral parietal, 8, left medial parietal, 9, left occipital and 10, left subcortical, 11, right lateral frontal, 12, right medial frontal, 13, right cingulate, 14, right insula, 15, right lateral temporal, 16, right medial temporal, 17, right lateral parietal, 18, right medial parietal, 19, right occipital, and 20, right subcortical) and a graphical representation of region connectivity summary where regions with above thresholds are indicated with a filled circle. On the right, significant negative correlations between S disp and sSneg are summarized in the same way.

by decreased pos strength with simultaneously increased neg strength were located in the orbito-frontal region, mesial temporal and inferior temporal region. This strength pattern indicates that the BOLD signal in these regions was anti-correlated to that of the majority of other regions. Interestingly, the other activity states or clusters showed very similar maxima and minima even though the strength dispersion during those phases was far less prominent. This can be interpreted as evidence that paroxysmal hypersynchrony enhances existing strength patterns in the brain rather than re-configuring them. Only about 45% of the CN (14 amyloid neg CN, seven amyloid pos CN) displayed cluster 8 activity in their task-free fMRI. The absence of hyperconnectivity phases in the other amyloid pos CN does not allow for the conclusion that they are free of such episodes though. It is possible that hyperconnectivity phases are less frequent, less severe or less widespread in these subjects and therefore not detected during the 8 min task-free fMRI with the approach used in this study. That being said it is equally possible that the hyperconnectivity phases indeed occur only in a subset of the amyloid pos and neg subjects who share a common unknown predisposition that renders the brain susceptible to paroxysmal hypersynchrony and that increasing brain amyloid enhances that predisposition. It will be necessary to investigate this question in longitudinal studies that acquire task-free fMRI over a longer time, e.g., 20–30 min.

The cognitive performance of participants in this study was within the age appropriate range and not different between subjects with and without increased brain amyloid. Nonetheless memory performance was negatively correlated with the duration of these paroxysmal hypersynchrony phases. This suggests that these phases might negatively affect cognitive function if they become longer or occur more frequently. To better understand how these transient hypersynchrony states interfere with memory performance, their spatial and temporal pattern in the mesial and lateral temporal lobe was investigated in more detail. Compared to other activity clusters, the high strength dispersion zone during cluster 8 activity engaged a larger part of the lateral temporal region (50% compared to 24–44% in other clusters) which led to a smaller moderate strength dispersion zone (23% compared to 35–55% in other clusters). These findings can be interpreted as evidence that the lateral temporal lobe is able to engage in the hypersynchronous activity causing the extreme strength dispersion in other parts of the brain. This was not the case in the mesial temporal region where the high strength dispersion zone was smaller (9% compared to 13–25% in other clusters) but the intermediate strength dispersion zone was larger than that of other activity clusters (55 vs. 31–49%). The increased zone of intermediate strength dispersion indicates that the hypersynchronous activity that dominates other brain regions is not able to completely overcome the anti-correlated BOLD activity in the mesial temporal region resulting in a functional disconnection of these structures during these phases. The finding of a negative correlation between memory performance and the frequency of windows in which extreme manifestations of this mesial temporal disconnection and of the temporal lateral hypersynchronization were observed also supports this hypothesis.

The extreme strength profile of cluster 8 activity, i.e., synchronization of the BOLD fluctuations over a large brain region but failure to engage mesial temporal lobe structures, can be explained by the findings of the structural connectivity analyses. While age had the expected effect on white matter connectivity, i.e., was negatively correlated with white matter connectivity in mostly frontal regions, amyloid load was positively correlated with white matter connectivity in lateral temporal lobe structures and to a lesser degree also in mesial temporal structures and in the precuneus. Although one has to be careful when interpreting DTI findings in regard of white matter integrity/functionality (Jones et al., 2013), the positive correlation with SUVR and the normal cognitive function suggests that amyloid did not have a negative impact on white matter connectivity at this early stage. This is in accordance with other cross-sectional studies that found normal or increased FA in amyloid positive CN in the absence of widespread tau accumulation (Racine et al., 2014; Wolf et al., 2015; Rieckmann et al., 2016; Kantarci et al., 2017). Normal or even increased white matter connectivity facilitates the generation and spread of the hypersynchronous activity hypothesized to be responsible for the prominent strength dispersion characterizing cluster 8 activity. It also supports normal between-region interactions as evidenced in the cross-modality analysis that found one or more regions with above threshold correlations between strength dispersion and white matter connectivity in almost all clusters. In contrast, amyloid had a negative effect on gray matter connectivity in the mesial temporal region as evidenced by the positive correlation between SUVR and atrophy-related connectivity in the right fusiform cortex that was accompanied by more widespread mesial temporal connectivity loss at more liberal statistical thresholds. It seems reasonable to assume that subtle atrophy in the mesial temporal region contributed to the functional disconnection of this region during cluster 8 activity. This assumption is also supported by the findings of the cross-modality analysis. Clusters 8 and 7 that are both characterized by a high strength dispersion show several regions with an above threshold number of negative correlations between atrophy affected gray matter connectivity and strength dispersion.

A SUVR exceeding the threshold used for amyloid positivity in this study indicates a diffuse, widespread pathology with amyloid deposits in the lateral superior temporal, lateral and midline frontal and parietal regions and beyond (Braak and Braak, 1997). The prolonged widespread hypersynchrony phases with maxima in the lateral temporal and temporoparietal, lateral frontal regions, and precuneus observed in the amyloid pos CN in this study are consistent with this widespread amyloid pathology. The circumscribed gray matter atrophy and the localized functional disconnection in the mesial temporal region however seem at odds with a diffuse pathology and also with the widely acknowledged observation that amyloid pathology correlates poorly with neurodegeneration. A circumscribed pathology with signs of neurodegeneration is commonly associated with tau pathology. Interestingly, the strength dispersion minima in the mesial, inferior temporal

and orbito-frontal regions are not only different from that seen in young controls (cf. **Supplementary Figure 1**) but also correspond well to the pattern of tau pathology at this stage (Schöll et al., 2016; Pontecorvo et al., 2017; Schultz et al., 2017; Sepulcre et al., 2017). The “dual” association of the mesial temporal structural and functional findings with amyloid (significant correlation with SUVR) on the one hand and with tau (regional preference, neurodegeneration) on the other hand is interesting because it ties into the observation that amyloid facilitates the development of widespread tau pathology that characterizes the clinical manifest stages of AD (Musiek and Holtzman, 2015). The mechanisms of this interaction are still far from clear and are one of the major research topics of the AD field (Pooler et al., 2015; Lewis and Dickson, 2016; Ayers et al., 2017). In the context of this study’s findings, it is particularly interesting that neuronal activity is supposed to play a major role in some of the proposed mechanisms. It is tempting to speculate that the type of widespread low level hypersynchronous activity that characterizes cluster 8 could represent a type of neuronal activity that is particularly well-suited to enable tau spreading. If this is true, amyloid pos subjects who show prolonged phases of this activity in task-free fMRI, would be expected to be at higher risk to develop a widespread tau pathology, cognitive impairment and brain atrophy associated with it than amyloid pos subjects who do not show hyperconnectivity phases or only very short phases. Given the negative correlation between atrophy-related gray matter connectivity and functional strength dispersion, it would be expected that the spreading tau leads to the development of widespread gray matter atrophy with secondary impact on white matter connectivity. As a consequence of the increasingly impaired structural connectivity the widespread functional hyperconnectivity of the early stage would be gradually replaced by a widespread hypoconnectivity in the later stages (Schultz et al., 2017). Identifying amyloid pos CN with hypersynchronous phases at an early stage and suppressing that activity with a suitable medication (Bakker et al., 2015) could prevent this hypothesized interaction between amyloid and tau and thus eventually delay or even prevent the development of cognitive impairment.

To our knowledge, this is the first study that combines a dynamic task-free fMRI analysis with white and gray matter connectivity analyses to investigate structure-function associations in amyloid pos and neg CN. There are previous studies that used different approaches of dynamic task-free analysis to investigate functional dynamics in preclinical and early AD. For example, Jones et al. (2012) investigated the dwell time of subnetworks in the dorsal and posterior DMN in AD patients and found shorter dwell times in brain states with posterior DMN contributions and longer dwell times in those with dorsal DMN contributions in AD compared to controls. Demirtaş et al. (2017) used effective connectivity to study global and regional fluctuations of synchronization over the whole range of the AD from healthy controls to fully developed AD and found a monotonous decrease over the disease course. Kang et al. (2017) used regional homogeneity (ReHo) to investigate functional synchronization in amyloid neg and

pos CN and found positive correlations between amyloid load and ReHo in the lingual gyrus, left fusiform gyrus, and right middle temporal gyrus in amyloid pos subjects, i.e., evidence for a localized hypersynchrony. Quevenco et al. (2017) finally used a sliding windows approach combined with PCA in CN and found a reduced anterior-posterior connectivity in CN whose cognitive functions worsened over 2 years compared to those who did not decline but no significant association between this connectivity reduction and amyloid load. While these studies clearly show that dynamic task-free fMRI analyses help to better understand the impact of beta amyloid on brain function, the study populations and/or analysis methods differ from the approach used in this study which complicates a comparison of the findings. The same is also true for the only previous study that investigated gray matter connectivity at the single subject level in amyloid pos and neg CN so far. Tijms et al. used a high resolution parcellation to assess similarities of gray matter structure instead of gray matter loss as was done in this study. They used graph analysis designed to look at sparse binary networks to describe gray matter disruptions and found a lower whole brain connectivity density and a less efficient network organization in amyloid pos CN (Tijms et al., 2016).

The study has several limitations. (1) Amyloid load was assessed using a region of interest approach to calculate global SUVR. Given the objective of this study it would have been desirable to use quantitative fluorbetapir maps. However, differences in the acquisition of the PET data at the two sites prevented the reconstruction of quantitative maps. (2) Given the potential association of some of the findings with tau pathology, it would have been desirable to obtain tau PET images as well. However, this was not possible due to budgetary restraints and the limited availability of the tracer at the time of this project. (3) The cross-modality analysis used simple spearman correlations between each and every roi of the two modalities and FDR to correct for multiple comparisons. This approach is not uncommon, but there exist more sophisticated multivariate statistical approaches, e.g., partial least square or sparse canonical correlation, and it cannot be excluded that these would have detected additional interesting associations between modalities. (4) The study population in this study was small and thus these findings have to be considered as preliminary and need to be confirmed in different and ideally larger populations that have amyloid and tau imaging, e.g., the new ADNI.

AUTHOR CONTRIBUTIONS

SM: Developed hypotheses and analysis methods, performed analysis, manuscript editing and writing; MW: Critical discussion of hypotheses and findings, manuscript editing and writing.

ACKNOWLEDGMENTS

The study was supported by a NIH grant R01 AG010897 to MW and a UCSF grant REAC/CTSI 37785–525205 to SM.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fnagi.2017.00418/full#supplementary-material>

Supplementary Figure 1 | Summary of cluster characterization in 20 young (mean age: 25.6 ± 4.8 , range 21–40) healthy and cognitively normal subjects who were studied with the same magnet and with the same imaging protocol. The task-free data was processed in the same way as described for the older subjects with the exception that the Spos and Sneg z-scores were calculated with all 20 subjects as reference. The cluster analysis identified 8 clusters (4 motion clusters),

the remaining clusters are displayed. The upper row shows the distribution of Spos in warm colors (please see color bar at the bottom of the figure), the middle row shows the Sneg distribution in cold colors (please see color bar at the bottom) and the lower row the maxima (>75 percentile) of S disp in red and the minima (<25 percentile) in blue. Please note that the Sdisp minima in the temporal lobes of these young subjects are confined to the parahippocampus and that the maxima extend into the inferior temporal lobe gyri. Finally, even though cluster 4 has a non-balanced positive profile, the dispersion is smaller than that of cluster 8 in the old population.

Supplementary Table 1 | Correlations No of cluster counts/subject with amyloid load and memory.

REFERENCES

- Ayers, J. I., Giasson, B. I., and Borchelt, D. R. (2017). Prion-like spreading in tauopathies. *Biol. Psychiatry*. doi: 10.1016/j.biopsych.2017.04.003. [Epub ahead of print].
- Bakker, A., Albert, M. S., Krauss, G., Speck, C. L., and Gallagher, M. (2015). Response of the medial temporal lobe network in amnesic mild cognitive impairment to therapeutic intervention assessed by fMRI and memory task performance. *Neuroimage Clin.* 7, 688–698. doi: 10.1016/j.nicl.2015.02.009
- Born, H. A. (2015). Neuroscience forefront review: seizures in Alzheimer's disease. *Neuroscience* 286, 251–263. doi: 10.1016/j.neuroscience.2014.11.051
- Braak, H., and Braak, E. (1997). Frequency of stages of Alzheimer related-lesions in different age categories. *Neurobiol. Aging* 18, 351–357. doi: 10.1016/S0197-4580(97)00056-0
- Brier, M. R., Thomas, J. B., Fagan, A. M., Hassenstab, J., Holtzman, D. M., Benzinger, T. L., et al. (2014). Functional connectivity and graph theory in preclinical Alzheimer's disease. *Neurobiol. Aging* 35, 757–768. doi: 10.1016/j.neurobiolaging.2013.10.081
- Busche, A. M., and Konnerth, A. (2016). Impairment of neuronal circuit function in Alzheimer's disease. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 371:201504429. doi: 10.1098/rstb.2015.0429
- Busche, M. A., Chen, X., Henning, H. A., Reichwald, J., Staufenbiel, M., Sakman, B., et al. (2012). Critical role of soluble amyloid beta for early hippocampal hyperactivity in a mouse model of Alzheimer's disease. *Proc. Natl. Acad. Sci.* 109, 8740–8745. doi: 10.1073/pnas.1206171109
- Demirtaş, M., Falcon, C., Tucholka, A., Gispert, J. D., Molinuevo, J. L., and Deco, G. (2017). A whole-brain computational modeling approach to explain the alterations in resting-state functional connectivity during progression of Alzheimer's disease. *Neuroimage Clin.* 16, 343–354. doi: 10.1016/j.nicl.2017.08.006
- Elman, J. A., Madison, C. M., Baker, S. L., Vogel, J. W., Mark, S. M., Crowley, S., et al. (2016). Effects of beta-amyloid on resting state functional connectivity within and between networks reflects known patterns of regional vulnerability. *Cereb. Cortex* 26, 695–707. doi: 10.1093/cercor/bhu259
- Grienberger, C., Rochefort, N. L., Adelsberger, H., Henning, H. A., Hill, D. N., Reichwald, J., et al. (2012). Staged decline of neuronal function in vivo in an animal model of Alzheimer's disease. *Nat. Commun.* 3:774. doi: 10.1038/ncomms1783
- Hutchison, R. M., Wormelsdorf, T., Allen, E. A., Bandettini, P. A., Calhoun, V. D., Corbetta, M., et al. (2013). Dynamic functional connectivity: promise, issues and interpretations. *Neuroimage* 80, 360–378. doi: 10.1016/j.neuroimage.2013.05.079
- Irfanoglu, M. O., Walker, L., Sarlls, J., Marceno, S., and Pierpaoli, C. (2012). Effects of image distortions originating from susceptibility variations and concomitant fields on diffusion MRI tractography results. *Neuroimage* 61, 275–288. doi: 10.1016/j.neuroimage.2012.02.054
- Jack, C. R., Wise, H. J., Weigand, S. D., Knopman, D. S., Lowe, V., Vermuri, P., et al. (2013). Amyloid-first and neurodegeneration-first profiles characterize incident amyloid PET positivity. *Neurology* 81, 1732–1740. doi: 10.1212/01.wnl.0000435556.21319.e4
- Jiang, Y., Huang, H., Abner, E., Broster, L. S., Jicha, G. A., Schmitt, F. A., et al. (2016). Alzheimer's biomarkers are correlated with brain connectivity in older adults differentially during resting and task states. *Front. Aging Neurosci.* 8:15. doi: 10.3389/fnagi.2016.00015
- Joliet, M., Jobard, G., Naveau, M., Delcroix, N., Petit, L., Zago, L., et al. (2015). AICHA: an atlas of intrinsic connectivity of homotopic areas. *J. Neurosci. Methods* 254, 46–59. doi: 10.1016/j.jneumeth.2015.07.013
- Jones, D. K., Knösche, T. R., and Turner, R. (2013). White matter integrity, fiber count and other fallacies: the do's and don'ts of diffusion MRI. *Neuroimage* 73, 239–254. doi: 10.1016/j.neuroimage.2012.06.081
- Jones, D. T., Knopman, D. S., Gunter, J. L., Graff-Radford, J., Vermuri, P., Bove, B. F., et al. (2016). Cascading network failure across Alzheimer's disease spectrum. *Brain* 139, 547–562. doi: 10.1093/brain/awv338
- Jones, D. T., Vermuri, P., Murphy, M. C., Gunter, J. L., Senjem, M. L., Machulda, M. M., et al. (2012). Non-stationarity in the resting brain's modular architecture. *PLoS ONE* 7:e39731. doi: 10.1371/journal.pone.0039731
- Kang, D. W., Choi, W. H., Jung, W. S., Um, Y. H., Chang, U. L., and Lim, H. K. (2017). Impact of amyloid burden on regional functional synchronization in the cognitively normal older adults. *Sci. Rep.* 7:14690. doi: 10.1038/s41598-017-15001-8
- Kantarci, K., Murray, M. E., Schwarz, C. G., Reid, R. I., Przybelski, S. A., Lesnick, T., et al. (2017). White-matter integrity on DTI and the pathologic staging of Alzheimer's disease. *Neurobiol. Aging* 56, 172–179. doi: 10.1016/j.neurobiolaging.2017.04.024
- Kellner, V., Menkes-Caspi, N., Beker, S., and Stern, E. A. (2014). Amyloid beta alters ongoing neuronal activity and excitability in the frontal cortex. *Neurobiol. Aging* 35, 1982–1991. doi: 10.1016/j.neurobiolaging.2014.04.001
- Kleen, J. K., Scott, R. C., Holmes, G. L., Roberts, D. W., Rundle, M. M., Testorf, M., et al. (2013). Hippocampal interictal epileptiform activity disrupts cognition in humans. *Neurology* 81, 18–24. doi: 10.1212/WNL.0b013e318297ee50
- Kramberger, M. G., Kareholt, I., Andersson, T., Winblad, B., Eriksdotter, M., and Jelic, V. (2013). Association between EEG Abnormalities and CSF biomarkers in a memory clinic cohort. *Dement. Geriatr. Cogn. Disord.* 36, 319–328. doi: 10.1159/000351677
- Leemans, A., Jeurissen, B., Sijbers, J., and Jones, D. K. (2009). “ExploreDTI: a graphical toolbox for processing, analyzing and visualizing diffusion MR data,” in *17th Annual Meeting of Intl Soc Mag Reson Med* (Hawaii), 3537.
- Leemans, A., and Jones, D. K. (2009). The B-matrix must be rotated when correcting for subject motion in DTI data. *Magn. Reson. Med.* 61, 1336–1349. doi: 10.1002/mrm.21890
- Lewis, J., and Dickson, D. W. (2016). Propagation of tau pathology: hypotheses, discoveries, an yet unresolved questions from experimental and human brain studies. *Acta Neurobiol.* 131, 27–48. doi: 10.1007/s00401-015-1507-z
- Liedorp, M., Stam, C. J., van der Flier, W. M., Pijnenburg, Y. A., and Scheltens, P. (2010). Prevalence and clinical significance of epileptiform EEG discharges in a large memory clinic cohort. *Dement. Geriatr. Cogn. Disord.* 29, 432–437. doi: 10.1159/000278620
- Liedorp, M., van der Flier, W. M., Hoogervorst, E. L. J., Scheltens, P., and Stam, C. J. (2009). Associations between patterns of EEG abnormalities and diagnosis in a large memory cohort. *Dement. Geriatr. Cogn. Disord.* 27, 18–23. doi: 10.1159/000182422
- Lim, H. K., Nebes, R., Snitz, B., Cohen, A., Mathis, C., Price, J., et al. (2014). Regional amyloid burden and intrinsic connectivity networks in cognitively normal elderly subjects. *Brain* 137, 3327–3338. doi: 10.1093/brain/awu271
- Matura, S., Prvulovic, D., Butz, M., Hartmann, D., Sepanski, B., Linnemann, K., et al. (2014). Recognition memory is associated with altered resting-state

- functional connectivity in people at genetic risk for Alzheimer's disease. *Eur. J. Neurosci.* 40, 3128–3135. doi: 10.1111/ejn.12659
- Mormino, E., Smiljic, A., Hyenga, A. O., Onami, S. H., Greicius, M. D., Rabinovici, G. D., et al. (2011). Relationships between beta-amyloid and functional connectivity in different components of the default mode network in aging. *Cereb. Cortex* 21, 2399–2407. doi: 10.1093/cercor/bhr025
- Mucke, L., and Selkoe, D. L. (2012). Neurotoxicity of amyloid β protein: synaptic and network dysfunction. *Cold Spring Harb. Perspect. Med.* 2:a006338. doi: 10.1101/cshperspect.a006338
- Mueller, S. G., Chao, L. L., Berman, B., and Weiner, M. W. (2011). Evidence for functional specialization of hippocampal subfields detected by MR subfield volumetry on high resolution images at 4T. *Neuroimage* 56, 851–857. doi: 10.1016/j.neuroimage.2011.03.028
- Murphy, K., Birn, E. M., Handwerker, D. A., Jones, T. B., and Bandettini, P. A. (2009). The impact of global signal regression on resting state correlations: are anti-correlated networks introduced? *Neuroimage* 44, 893–905. doi: 10.1016/j.neuroimage.2008.09.036
- Musiek, E. S., and Holtzman, D. M. (2015). Three dimensions of the amyloid hypothesis: time, space and “wingman.” *Nat. Neurosci.* 18, 800–806. doi: 10.1038/nn.4018
- Palop, J. J., Chin, J., Roberson, E. D., Wang, J., Thwin, M. T., Bien-Ly, N., et al. (2007). Aberrant excitatory neuronal activity and compensatory remodeling of inhibitory hippocampal circuits in mouse models of Alzheimer's disease. *Neuron* 55, 697–711. doi: 10.1016/j.neuron.2007.07.025
- Palop, J. J., and Mucke, L. (2010). Amyloid-beta-induced neuronal dysfunction in Alzheimer's disease: from synapses toward neural networks. *Nat. Neurosci.* 13, 812–818. doi: 10.1038/nn.2583
- Palop, J. J., and Mucke, L. (2016). Network abnormalities and interneuron dysfunction in Alzheimer disease. *Nat. Rev. Neurosci.* 17, 777–792. doi: 10.1038/nrn.2016.141
- Pontecorvo, M. J., Devous, M. D., Navitsky, M., Lu, M., Salloway, S., Schaerf, F. W., et al. (2017). Relationships between flortaucipir PET binding and amyloid burden, clinical diagnosis and cognition. *Brain* 140, 748–763. doi: 10.1093/brain/aww334
- Pooler, A. M., Polydoro, M., Maury, E. A., Nicholls, S. B., Reddy, S. M., Wegmann, S., et al. (2015). Amyloid accelerates tau propagation and toxicity in a model of early Alzheimer's disease. *Acta Neuropathol. Commun.* 3:14. doi: 10.1186/s40478-015-0199-x
- Power, J. D., Barnes, K. A., Snyder, A. Z., Schlaggar, B. L., and Petersen, S. E. (2012). Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *Neuroimage* 59, 2142–2154. doi: 10.1016/j.neuroimage.2011.10.018
- Quevenco, F., Preti, M. G., van Bergen, J. M. G., Hua, J., Wyss, M., Li, X., et al. (2017). Memory performance-related dynamic brain connectivity indicates pathological burden and genetic risk for Alzheimer's disease. *Alzheimer Res. Ther.* 9:24. doi: 10.1186/s13195-017-0249-7
- Racine, A. M., Adluru, N., Alexander, A. L., Christian, B. T., Okonkwo, O. C., Oh, J., et al. (2014). Associations between white matter microstructure and amyloid burden in preclinical Alzheimer's disease. a multimodal imaging investigation. *Neuroimage Clin.* 4, 604–616. doi: 10.1016/j.nicl.2014.02.001
- Rieckmann, A., Van Dijk, K. R. A., Sperling, R. A., Johnson, K. A., Buckner, R. L., and Hedden, T. (2016). Accelerated decline in white matter integrity in clinically normal individuals at risk for Alzheimer's disease. *Neurobiol. Aging* 42, 177–188. doi: 10.1016/j.neurobiolaging.2016.03.016
- Rubinov, M., and Sporns, O. (2011). Weight conserving characterization of complex functional brain networks. *Neuroimage* 56, 2068–2079. doi: 10.1016/j.neuroimage.2011.03.069
- Schöll, M., Lockhart, S. N., Schonhaut, D. R., O'Neil, J. P., Janabi, M., Ossenhoppe, R., et al. (2016). PET imaging of tau disposition in the aging human brain. *Neuron* 89, 971–982. doi: 10.1016/j.neuron.2016.01.028
- Schultz, A. P., Chhatwal, J. P., Hedden, T., Mormino, E., Hanseeuw, B. J., Sepulcre, J., et al. (2017). Phases of hyperconnectivity and hypoconnectivity in the default mode and salience networks track with amyloid and tau in clinically normal individuals. *J. Neurosci.* 37, 4323–4331. doi: 10.1523/JNEUROSCI.3263-16.2017
- Sepulcre, J., Sabuncu, M. R., Li, Q., El Fakhri, G., and Sperling, R. (2017). Tau and amyloid-beta proteins distinctively associate to functional network changes in the aging brain. *Alzheimers Dement.* 13, 1261–1269. doi: 10.1016/j.jalz.2017.02.011
- Shah, D., Praet, J., Hernandez, A. L., Höfling, C., Anckaerts, C., Bard, F., et al. (2016). Early pathogenic amyloid induces hypersynchrony of BOLD resting state networks in transgenic mice and provides an early therapeutic window before amyloid plaque deposition. *Alzheimers Dement.* 12, 964–976. doi: 10.1016/j.jalz.2016.03.010
- Sheline, Y. I., Raichle, M. E., Snyder, A. Z., Morris, J. C., Head, D., Wang, S., et al. (2010). Amyloid plaques disrupt resting state default mode network connectivity in cognitively normal elderly. *Biol. Psychiatry* 67, 584–587. doi: 10.1016/j.biopsych.2009.08.024
- Smits, L. L., Liedorp, M., Koene, T., Roos-Reuling, I. E., Lemstra, A. W., Scheltens, P., et al. (2011). EEG abnormalities are associated with different cognitive profiles in Alzheimer's disease. *Dement. Geriatr. Cogn. Disord.* 31, 1–6. doi: 10.1159/000322202
- Sperling, R. A., LaViolette, P. S., O'Keefe, K., O'Brien, J., Rentz, D. M., Pihlajamaki, M., et al. (2009). Amyloid deposition is associated with impaired default network function in older persons without dementia. *Neuron* 63, 178–188. doi: 10.1016/j.neuron.2009.07.003
- Stargardt, A., Swaab, D. F., and Bossers, K. (2015). The storm before the quiet: neuronal hyperactivity and Abeta in the pre-symptomatic stages of Alzheimer's disease. *Neurobiol. Aging* 36, 1–11. doi: 10.1016/j.neurobiolaging.2014.08.014
- Steininger, S. C., Liu, X., Gietl, A., Wyss, M., Schreiner, S., Gruber, E., et al. (2014). Cortical amyloid beta in cognitively normal elderly adults is associated with decreased network efficiency within the cerebro-cerebellar system. *Front. Aging Neurosci.* 6:52 doi: 10.3389/fnagi.2014.00052
- Talantova, M., Sanz-Blasco, S., Zhang, X., Xia, P., Akhtar, M. W., Okamoto, S. I., et al. (2013). Abeta induces astrocytic glutamate release, extrasynaptic NMDA receptor activation and synaptic loss. *Proc. Natl. Acad. Sci.* 110, 2518–2527. doi: 10.1073/pnas.1306832110
- Tijms, B. M., ten Kate, M., Wink, A. M., Visser, P. J., Ecay, M., Clerique, M., et al. (2016). Gray matter network disruptions and amyloid beta in cognitively normal adults. *Neurobiol. Aging* 37, 154–160. doi: 10.1016/j.neurobiolaging.2015.10.015
- Vossel, K. A., Beagle, A. J., Rabinovici, G. D., Shu, H., Lee, S. E., Naasan, G., et al. (2013). Seizures and epileptiform activity in the early stages of Alzheimer's disease. *JAMA Neurol.* 70, 1158–1166. doi: 10.1001/jamaneurol.2013.136
- Wang, L., Brier, M. R., Snyder, A. Z., Thomas, J. B., Fagan, A. M., Xiong, C., et al. (2013). Cerebrospinal fluid Abeta42, phosphorylates Tau181, and resting-state functional connectivity. *JAMA Neurol.* 70, 1242–1248. doi: 10.1001/jamaneurol.2013.3253
- Whitfield-Gabrieli, S., and Nieto-Castagnon, A. (2012). Conn: a functional connectivity toolbox for correlated and anticorrelated brain networks. *Brain Connect.* 2, 125–141. doi: 10.1089/brain.2012.0073
- Wolf, D., Fischer, F. U., Scheurich, A., and Fellgiebel, A. (2015). Non linear association between cerebral amyloid deposition and white matter microstructure in cognitively healthy older adults. *J. Alzheimers Dis.* 47, 117–127. doi: 10.3233/JAD-150049

Conflict of Interest Statement: MW has been on scientific advisory boards for Pfizer and BOLT International; has been a consultant for Pfizer Inc., Janssen, KLI Associates, Easton Associates, in Thought, INC Research, Inc., Alzheimer's Drug Discovery Foundation and Sanofi-Aventis Groupe; has received funding for travel from Pfizer, Novartis, Tohoku University, MCI Group, France, Travel eDreams, Inc., Neuroscience School of Advanced Studies (NSAS), Danone Trading, BV, CTAD ANT Congres; has received honoraria from Pfizer, Tohoku University, and Danone Trading, BV; has research support from Merck and Avid.

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

Copyright © 2017 Mueller and Weiner. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Disrupted Thalamus White Matter Anatomy and Posterior Default Mode Network Effective Connectivity in Amnestic Mild Cognitive Impairment

Thomas Alderson^{1*}, Elizabeth Kehoe², Liam Maguire¹, Dervla Farrell², Brian Lawlor³, Rose A. Kenny³, Declan Lyons⁴, Arun L. W. Bokde² and Damien Coyle¹

¹Intelligent Systems Research Centre, University Ulster, Derry, United Kingdom, ²Trinity College Institute of Neuroscience and Cognitive Systems Group, Discipline of Psychiatry, School of Medicine, Trinity College Dublin, Dublin, Ireland, ³Mercer's Institute for Research on Ageing, St. James's Hospital, Trinity College Institute of Neuroscience, Trinity College Dublin, Dublin, Ireland, ⁴St. Patrick's Hospital, Dublin, Ireland

Alzheimer's disease (AD) and its prodromal state amnestic mild cognitive impairment (aMCI) are characterized by widespread abnormalities in inter-areal white matter fiber pathways and parallel disruption of default mode network (DMN) resting state functional and effective connectivity. In healthy subjects, DMN and task positive network interaction are modulated by the thalamus suggesting that abnormal task-based DMN deactivation in aMCI may be a consequence of impaired thalamo-cortical white matter circuitry. Thus, this article uses a multimodal approach to assess white matter integrity between thalamus and DMN components and associated effective connectivity in healthy controls (HCs) relative to aMCI patients. Twenty-six HC and 20 older adults with aMCI underwent structural, functional and diffusion MRI scanning using the high angular resolution diffusion-weighted acquisition protocol. The DMN of each subject was identified using independent component analysis (ICA) and resting state effective connectivity was calculated between thalamus and DMN nodes. White matter integrity changes between thalamus and DMN were investigated with constrained spherical deconvolution (CSD) tractography. Significant structural deficits in thalamic white matter projection fibers to posterior DMN components posterior cingulate cortex (PCC) and lateral inferior parietal lobe (IPL) were identified together with significantly reduced effective connectivity from left thalamus to left IPL. Crucially, impaired thalamo-cortical white matter circuitry correlated with memory performance. Disrupted thalamo-cortical structure was accompanied by significant reductions in IPL and PCC cortico-cortical effective connectivity. No structural deficits were found between DMN nodes. Abnormal posterior DMN activity may be driven by changes in thalamic white matter connectivity; a view supported by the close anatomical and functional association of thalamic nuclei effected by AD pathology and the posterior DMN nodes. We conclude that dysfunctional posterior DMN activity in aMCI is consistent with disrupted cortico-thalamo-cortical processing and thalamic-based dissemination of hippocampal disease agents to cortical hubs.

Keywords: diffusion MRI, tractography, effective connectivity, Alzheimer's disease, mild cognitive impairment, default mode network, thalamus, resting state

OPEN ACCESS

Edited by:

Stefan Teipel,
German Center for
Neurodegenerative Diseases (HZ),
Germany

Reviewed by:

Marco Duering,
Klinikum der Universität München,
Germany
Juan Zhou,
Duke-NUS Medical School,
Singapore

*Correspondence:

Thomas Alderson
thomashenryalderson@gmail.com

Received: 17 July 2017

Accepted: 26 October 2017

Published: 08 November 2017

Citation:

Alderson T, Kehoe E, Maguire L,
Farrell D, Lawlor B, Kenny RA,
Lyons D, Bokde ALW and Coyle D
(2017) Disrupted Thalamus White
Matter Anatomy and Posterior
Default Mode Network Effective
Connectivity in Amnestic Mild
Cognitive Impairment.
Front. Aging Neurosci. 9:370.
doi: 10.3389/fnagi.2017.00370

INTRODUCTION

Alzheimer's disease (AD) is a chronic neurodegenerative disorder affecting approximately 6% of people over the age of 65 and accounting for 60%–70% of dementia cases (Burns and Iliffe, 2009). Typically, the AD-prodromal stage presents as mild cognitive impairment (MCI; Stephan et al., 2012) clinically defined as cognitive difficulties beyond those expected based on age and education, but insufficient to interfere with daily activities (Petersen et al., 1999; Petersen, 2004). MCI can present with a variety of symptoms but is termed amnesic MCI (aMCI) in cases where memory loss is the predominant symptom.

In AD, the first neurofibrillary tangles appear in the parahippocampal regions (Stage I) followed later, and accompanied by cognitive symptoms, in the hippocampus formation (stage III; Braak and Braak, 1991a,b, 1995). Understandably, this knowledge has reinforced focus on the hippocampus in the context of memory loss in AD but much less well-known and less well-understood are the appearance of tangles and plaques in the thalamic nuclei in parallel with those in the hippocampus. Their appearance is often characterized as an event downstream of the hippocampus pathology transmitted by the projections of the mammillary bodies, but this view is challenged by metabolic studies indicating that the earliest consistent declines occur not in hippocampus but in posterior cingulate cortex (PCC; Minoshima et al., 1994, 1997) where amyloid deposition is highest (Buckner et al., 2005; Mintun et al., 2006). The thalamus, with its dense network of reciprocal interconnections with both hippocampus and PCC, is therefore implicated by association (Vann et al., 2009; Aggleton et al., 2010).

Such a view is supported by detection of thalamic atrophy in pre-symptomatic familial AD on average 5.6 years prior to expected symptom onset (Ryan et al., 2013) together with increased amyloid burden (Knight et al., 2011a,b) and substantial evidence suggesting that thalamic atrophy is present in MCI prior to AD (Chételat et al., 2005; Shiino et al., 2006; de Jong et al., 2008; Ferrarini et al., 2008; Cherubini et al., 2010; Roh et al., 2011; Pedro et al., 2012; Zhang et al., 2013). Structural irregularities have a sufficient impact on thalamo-cortical circuits to allow healthy subjects to be differentiated from those with MCI through impaired functional integrity (Cantero et al., 2009). Conversely, carriers of the apolipoprotein $\epsilon 2$ allele i.e., those showing a genetic predisposition against developing AD, demonstrate significantly enhanced functional (Patel et al., 2013) and structural (Chiang et al., 2012) integrity of the thalamus.

Analysis of low frequency BOLD signal oscillations have revealed several resting state networks. Of these, the default mode network (DMN; Raichle et al., 2001; Greicius et al., 2003; Damoiseaux et al., 2006) has consistently been identified as dysfunctional in both MCI and AD in the context of amyloid burden (Hedden et al., 2009; Drzezga et al., 2011; Mormino et al., 2011; Sheline et al., 2011) and genetic risk (Roses, 1996; Sheline et al., 2010; Wang et al., 2012; Chhatwal et al., 2013).

The DMN comprises medial prefrontal cortex (mPFC), middle temporal gyrus (MTG), lateral inferior parietal lobes (IPL), PCC and hippocampus regions. These nodes have been identified as important hubs within the cortex (Buckner et al., 2009) whose persistent background activity and dense, long range interconnectivity may facilitate the early deposition and prion-like transmission of amyloid plaques (Wermke et al., 2008; Raj et al., 2012). DMN topography is therefore recapitulated in the pattern of atrophy, hypometabolism and amyloid deposition within the cortex (Buckner et al., 2005, 2008).

Thalamus appears to play a role modulating distributed cortical networks (Di and Biswal, 2014). It is therefore of note, that direct structural connections between the thalamus and DMN (or thalamo-DMN pathway) components have been described *in vivo* using diffusion tensor imaging (DTI; Fernández-Espejo et al., 2012) and that these are sites of atrophy (Zarei et al., 2010). Crucially, lesions to the thalamus are known to cause DMN dysfunction (Jones et al., 2011). One suggestion is that abnormal task-induced deactivation of DMN response patterns in aMCI are a consequence of impaired thalamo-cortical signaling (Pihlajamäki and Sperling, 2009).

The thalamus sends widespread connections to its ipsilateral cortical hemisphere which are returned via cortico-thalamic feedback connections. Together these form a thalamo-cortico-thalamic feedback loop (Sherman and Guillery, 2006; Sherman, 2007; Zhang et al., 2008, 2010). Such an arrangement is critical for generating the ubiquitous oscillations of the cortex recorded by EEG and fMRI but its contribution (and other subcortical components) to regulating the DMN in health and disease is largely unexplored. On this basis, we chose to investigate the impact of impaired thalamo-cortical microscopic white matter anatomy on interactions in the DMN in aMCI patients.

We performed constrained spherical deconvolution (CSD)-based probabilistic fiber tractography of the thalamo-DMN white matter pathways in a cohort of older adults with aMCI and healthy age-matched controls. We also examined the effective connectivity of the resting state thalamo-DMN interactions using a spatio-temporal formulation of Granger Causality (GC). In contrast to simple statistical correlation (i.e., functional connectivity), effective connectivity is more ambitious and attempts to quantify the causal influence one region exerts over another. Given that thalamo-cortical neural signals appear to coordinate distributed networks (Di and Biswal, 2014), such an approach provides greater scope for clarifying the interactions between thalamus and cortex during the transition between health and disease. We predicted that abnormal DMN causal activity would be linked to structural deficits in the thalamo-DMN pathway.

MATERIALS AND METHODS

Participants

Twenty six HC participants and 20 older participants with aMCI took part in the study. The HCs were community-dwelling

older adults recruited from the greater Dublin area (Ireland) via newspaper advertisements. They underwent a health screening questionnaire and a neuropsychological assessment, the Consortium to Establish a Registry for Alzheimer's Disease (CERAD; Morris et al., 1989), in order to rule out possible cognitive impairment before inclusion in the study. The CERAD battery has been shown to be sensitive to the presence of age related cognitive decline (Welsh et al., 1991, 1992). All of the older participants included in the study scored no more than 1.5 SD below the standardized mean scores for subjects of a similar age and education level on any of the sub-tests. The aMCI participants were recruited from memory clinics in St. James Hospital and St. Patrick's Hospital in Dublin, Ireland, and were diagnosed by a clinician according to the Peterson criteria (Peterson et al., 1999)—i.e., abnormal memory scores for age and education level with no dementia. Four were single amnesic aMCI, and 16 were multi-domain aMCI (Peterson, 2004). Neuropsychological measures were administered or supervised by an experienced neuropsychologist and included the Mini-Mental State Examination (MMSE; Folstein et al., 1975) and Cambridge cognitive examination (Huppert et al., 1995).

All of the participants were right-handed with no history of head trauma, neurological disease, stroke, transient ischemic attack, heart attack, or psychiatric illness. They completed the Geriatric Depression Scale (GSD; Yesavage et al., 1983), the Eysenck Personality Questionnaire Revised Edition Short Scale (EPQ-R; Eysenck and Eysenck, 1994), and a Cognitive Reserve Questionnaire (Rami et al., 2011) before the MRI scan (Table 1). The groups did not differ in terms of age, gender, education level, or levels of cognitive reserve as assessed by the self-report Cognitive Reserve Questionnaire. The aMCI group had lower MMSE scores, higher GDS scores, and scored lower on the EPQ measure of extraversion than the HC group. The study had full ethical approval from the St. James Hospital and the Adelaide and Meath Hospital, incorporating the National Children's Hospital Research Ethics Committee and St. Patrick's University Hospital Research Ethics Committee. All subjects gave written informed consent in accordance with the Declaration of Helsinki.

TABLE 1 | Results of independent samples *t*-tests, except for gender which was compared with a Fischer's exact test.

| | HC (<i>n</i> = 26) | aMCI (<i>n</i> = 20) | <i>p</i> * (<i>df</i> = 44) |
|-----------|---------------------|-----------------------|------------------------------|
| Gender | 15 M, 11 F | 10 M, 10 F | 1.00 |
| Age | 69.30 ± 6.35 | 69.05 ± 7.55 | 0.90 |
| Ethnicity | White (Irish) | White (Irish) | - |
| Education | 13.38 ± 3.73 | 14.32 ± 3.02 | 0.38 |
| MMSE | 28.65 ± 0.85 | 27.05 ± 2.17 | 0.0013 |
| GDS | 0.77 ± 1.07 | 2.58 ± 2.27 | 0.0008 |
| EPQ E | 8.04 ± 2.47 | 5.53 ± 3.37 | 0.0061 |
| EPQ N | 2.69 ± 2.43 | 3.78 ± 3.39 | 0.21 |
| CR | 16.65 ± 3.62 | 16.58 ± 4.97 | 0.95 |

Standard deviations are indicated. Statistically significant differences are indicated in bold. MMSE, Mini-Mental State Exam; GDS, geriatric depression scale; EPQ E, Eysenck personality questionnaire extraversion scale; EPQ N, Eysenck personality questionnaire neuroticism scale; CR, cognitive reserve scale.

MRI Data Acquisition

Whole-brain high angular resolution diffusion imaging (HARDI) data were acquired on a 3.0 Tesla Philips Intera MR system (Best, Netherlands) equipped with an eight channel head coil. A parallel sensitivity encoding (SENSE) approach (Pruessmann et al., 1999) with a reduction factor of two was used during the diffusion weighted image (DWI) acquisition. Single-shot spin echo-planar imaging was used to acquire the DWI data with the following parameters: echo time (TE) = 79 ms, repetition time (TR) = 20,000 ms, field of view (FOV) = 248 mm, matrix = 112 × 112, isotropic voxel of 2.3 mm × 2.3 mm × 2.3 mm, and 65 slices with 2.3 mm thickness with no gap between the slices. Diffusion gradients were applied in 61 isotropically distributed orientations with $b = 3000 \text{ s/mm}^2$, and four images with $b = 0 \text{ s/mm}^2$ were also acquired. A high-resolution 3D T1-weighted anatomical image was acquired for each participant with the following parameters: TE = 3.9 ms, TR = 8.5 ms, FOV = 230 mm, slice thickness = 0.9 mm, voxel size = 0.9 mm × 0.9 mm × 0.9 mm. Resting-state fMRI data were also acquired during the scanning session. The scan lasted for 7 min during which time the participants were asked to keep their eyes open and fixate on a cross hairs in the center of a screen behind the MR scanner, visible via a mirror. The BOLD signal changes were measured using a T2*-weighted echo-planar imaging sequence with TE = 30 ms and TR = 2000 ms. Each volume of data covered the entire brain with 39 slices, and the slices were acquired in interleaved sequence from inferior to superior direction. Two-hundred and ten volumes of data were acquired, with voxel dimensions of 3.5 mm × 3.5 mm × 3.85 mm and a 0.35 mm gap between the slices.

Face-Name Encoding and Recognition Task Protocol

Relationships between the participants' structural/effective connectivity measures and memory were subsequently examined using data obtained from a face-name recognition task following the resting-state scan. The participants viewed a series of 27 emotional faces (Erwin et al., 1992) with a name presented underneath each one. This task was an implicit memory task, in that the participants later completed a surprise memory tasks to test their retention of both the faces and the face-name pairs, however, at the time of encoding, they were not explicitly asked to remember the face-name pairs. Rather, the participants were instructed to judge whether the names matched or suited the faces. It was explained that this was a subjective decision, with no right or wrong answer. The participants responded yes or no by pressing a button on a MR-compatible response pad held in their right or left hand, respectively, using the index finger of either hand. Each face-name combination was presented for 4 s and was shown twice during the run. The faces were positive, negative, or neutral in valence and there were equal numbers of valence types as well as gender. The presentation of the face-name pairs was grouped according to the emotional valence of the faces. In each instance, a group of either two, three, or four faces of one valence type was presented randomly using an event-related paradigm,

subsequently, there was a delay during which a white cross hair was presented (control condition). The duration of the white cross was varied according to the duration of the face stimulus. For instance, if a single face was presented for 4 s the subsequent white cross was also shown for 4 s and then the next block of faces began. The stimuli were delivered using Presentation v.16.1 (Neurobehavioral Systems, Albany, CA, USA).

Approximately 15 min following the encoding phase, the participants performed a short computer-based recognition task. The emotional faces were presented one at a time on a black background with three names underneath. One of the names was the correct name; one name was a name that had been paired with a different face (distractor; incorrect name), while the third name was a new name (foil; incorrect name). The participants responded by pressing a button on the left, middle, or right side of a keyboard to correspond with the relative position of the name on the screen. The stimuli were presented for 5 s and followed by an inter-trial interval of 5 s. This longer trial length was to facilitate performance of this task as it was quite challenging. Before the task began, the participants completed a short practice run of five trials.

Resting State Pre-Processing

fMRI data processing was carried out using FMRI Expert Analysis Tool (FEAT) Version 6.00, part of FMRIB's Software Library (FSL)¹. Registration to high resolution structural and/or standard space images was carried out using FNIRT (Andersson et al., 2007). The following pre-statistics processing was applied; motion correction using MCFLIRT (Jenkinson et al., 2002), slice-timing correction using Fourier-space time-series phase-shifting, non-brain removal using BET (Smith, 2002), spatial smoothing using a Gaussian kernel of FWHM 3.0 mm, grand-mean intensity normalization of the entire 4D dataset by a single multiplicative factor, highpass temporal filtering (Gaussian-weighted least-squares straight line fitting, with sigma = 50.0 s).

Resting State Effective Connectivity

GC is a standard statistical tool for detecting the directional influence one system component exerts over another. The concept, originally introduced by Wiener (1956), and later incorporated into a data analysis framework by Granger (1969) is described as follows. If historical information from time series X significantly improves prediction accuracy of the future of time series Y in a multivariate autoregressive model (MVAR), GC is identified. This may be viewed as a measure of model prediction error where GC quantifies the reduction in prediction error when past values of X are included in the explanatory variables of Y (Schelter et al., 2006). By fitting a time invariant MVAR model to the experimental time series the classic GC formulation ignores crucial time-varying properties of the system. Such an approach makes the tacit assumption that the longer the time series, the more reliable the GC estimates. While this may be correct in static circuit representations (Smith et al., 2011), under time-varying conditions this principle

is no longer valid. A more robust method is to divide the time series into equal windows and consider them separately. Here, an optimal trade-off between the length of the time windows and the accuracy of the estimated coefficients for each window must be determined. Time windows that are too short prevent the accurate estimation of parameters, while time windows that are too long increase the probability of incorrect inferences of GC. Accordingly, the current article utilizes a novel spatio-temporal GC formulation to quantify the effective connectivity changes between region of interest (ROI; Luo et al., 2013). In this framework, finding the optimal time window length reduces to the solution of a constrained optimization problem,

$$\min_{l_0(m)} \left(GC_{\text{err}}(l_0(m)) + \frac{1}{GC_{\text{avg}}(l_0(m))} \right)$$

where we seek to simultaneously minimize model prediction error GC_{err} (i.e., the weighted average of the variances of the residuals in each time window) and maximize detected causality information GC_{avg} (i.e., the average GC over all time windows). This is performed for time windows of different length $l_0(m) = t_1, \dots, t_m$. The time window producing the lowest Bayesian information criterion (BIC) is considered optimal. By considering optimal time windows, the spatio-temporal framework allows a more reliable and precise estimate of GC in experimental datasets with time varying properties. This approach has been shown to yield more accurate estimates of GC on resting state fMRI data than traditional GC metrics. In this case, the last 208 time points for each region under consideration were extracted from the functional image volume and divided into four windows with the first two time points removed to avoid start up transients. In terms of spatial resolution, GC is calculated between all pairs of voxels from the two ROI under consideration. The mean GC among all pairs of voxels was then used as the final estimate.

CSD White Matter Tractography Using MRtrix3

A method for controlling free water contamination of tissue and the resultant partial volume effects is especially important around the fornix where atrophy and cerebrospinal fluid (CSF) is prevalent. The free water elimination technique (Pasternak et al., 2009) has been successfully applied in previous tractography studies of ageing and aMCI (Metzler-Baddeley et al., 2012a,b; Fletcher et al., 2014; Kehoe et al., 2015) however at higher b-values the Gaussian assumption underlying the bi-tensor model is no longer valid and a more simple heuristic is indicated. Accordingly, we use the standard free water elimination approach to identify and mask voxels with high free water content but fit the conventional DTI model to each voxel.

The *dwipreproc* preprocessing script was to perform eddy current-induced distortion and motion correction using the FSL tool eddy (Andersson and Sotiropoulos, 2016). The standard MRtrix3 processing script *dwibiascorrect* was used to eliminate low frequency intensity inhomogeneities across the DWI series. The script uses bias field correction algorithms available in the FSL software package (Zhang et al., 2001).

¹www.fmrib.ox.ac.uk/fsl

Probabilistic white matter tractography was performed on the DWIs using the MRtrix3 software package². Crossing fibers were resolved using the CSD algorithm (Tournier et al., 2004, 2012). MRtrix3 pre-processing included computing the diffusion tensors images (or diffusion ellipsoids) for each voxel from which the fractional anisotropy (FA), axial (DA), radial (RD) and mean (MD) diffusivity.

Whole-brain tractography was performed using every voxel as a seed point. The principle diffusion orientation at each point was estimated by the CSD tractography algorithm, which propagated in 0.1 mm steps along this direction. At each new location the fiber orientation(s) was estimated before the tracking moved a further 0.1 mm along the direction that subtended the smallest angle to the current trajectory. A trajectory was followed through the data until the scaled height of the fiber orientation density function peak dropped below the default threshold, or the direction of the pathway changed through an angle of more than 90°.

Anatomical masks were used to divide the results into circumscribed regions. The DMN was defined by probabilistic template (Wang et al., 2014) and the hippocampus and thalamus using the Harvard-Oxford subcortical structural atlas (Figure 1). Streamlines beginning in one mask and terminating in another were considered in a pairwise fashion for all ROI. In addition, the FSL tool FMRIB's Automated Segmentation Tool (FAST) was used to derive a white matter brain mask to constrain tractography. Any tracks tracts exiting the white matter were considered spurious and discarded. Tracts were prevented from propagating between hemispheres by a stop region placed down the midline corresponding to the corpus callosum.

Statistically significant differences in the mean FA, DA, RD and MD of tracks in HC vs. aMCI were tested by way of a two tailed two sample *t*-test at $p < 0.05$ corrected for multiple comparisons.

Independent Component Analysis (ICA)

The DMN was identified for each subject using ICA. Analysis was carried out using Probabilistic ICA (Beckmann and Smith, 2004) as implemented in Multivariate Exploratory Linear Decomposition into Independent Components (MELODIC) Version 3.14, part of FSL. The following data pre-processing was applied to the input data: masking of non-brain voxels, voxel-wise de-meaning of the data, normalization of the voxel-wise variance. Pre-processed data were whitened and projected into a 62-dimensional subspace using probabilistic Principal Component Analysis where the number of dimensions was estimated using the Laplace approximation to the Bayesian evidence of the model order (Minka, 2001; Beckmann and Smith, 2004). The whitened observations were decomposed into sets of vectors which describe signal variation across the temporal domain (time-courses) and across the spatial domain (maps) by optimizing for non-Gaussian spatial source distributions using a fixed-point iteration technique (Hyvärinen, 1999). Estimated component maps were divided by the standard deviation of the residual noise and thresholded by

fitting a mixture model to the histogram of intensity values (Beckmann and Smith, 2004). The number of components was automatically estimated. The component corresponding to the DMN was selected by cross correlating all the components with a probabilistic DMN template (Wang et al., 2014). The fMRI BOLD signal was extracted from DMN components mPFC, MTG, IPL and PCC, combined with those extracted from hippocampus and thalamus masks, and analyzed using the spatio-temporal GC method to determine the effective connectivity. A standard two tailed *t*-test was used to determine significant differences between the HC and aMCI patients at $p < 0.05$ corrected for multiple comparisons.

RESULTS

Comparison of Resting State Thalamo-DMN Effective Connectivity in HC vs. aMCI Subjects

The spatio-temporal GC effective connectivity analysis revealed significant differences in a circumscribed set of regions at the Bonferroni corrected threshold of $p < 0.0014$.

In aMCI, several incoming connections to PCC and left IPL showed reduced casual connectivity. An especially pronounced decrease in causal interaction to left IPL from other DMN components, hippocampus, and thalamus was observed (Figure 2A). Reduced connectivity to left IPL included incoming connections from left thalamus ($t_{(44)} = 3.77, p < 0.001$), left ($t_{(44)} = 4.3, p < 0.0001$) and right ($t_{(44)} = 3.80, p < 0.001$) MTG and from right IPL ($t_{(44)} = 3.83, p < 0.001$). These changes correspond to a highly significant ($t_{(44)} = 5.10, p < 0.00001$) decrease in average FA in the white matter between left thalamus and left IPL (Figure 2B).

Also in aMCI, significant reductions in connectivity to PCC from left MTG ($t_{(44)} = 3.93, p < 0.001$) were found, together with significant reductions in connectivity to right IPL from PCC ($t_{(44)} = 3.73, p < 0.001$) (Figure 2A).

Comparison of Thalamo-DMN Microstructural Integrity in HC vs. aMCI Subjects

In aMCI, CSD white matter tractography identified statistically significant increases at the Bonferroni corrected threshold in average DA, RD and MD in the white matter fiber pathways connecting thalamus to hippocampus, PCC and IPL (Figure 3). Significant decreases in average FA were also detected in the white matter between thalamus and IPL. These included:

Significant decreases in FA (Figure 2B) between right thalamus and right IPL ($t_{(44)} = 3.81, p < 0.001$) and between left thalamus and left IPL ($t_{(44)} = 5.24, p < 0.00001$).

Significant increases in DA (Figure 2C) between right thalamus and right hippocampus ($t_{(44)} = -4.68, p < 0.0001$), right hippocampus and PCC ($t_{(44)} = -4.02, p < 0.001$), left thalamus and left hippocampus ($t_{(44)} = -5.33, p < 0.00001$), left

²<http://www.mrtrix.org/>

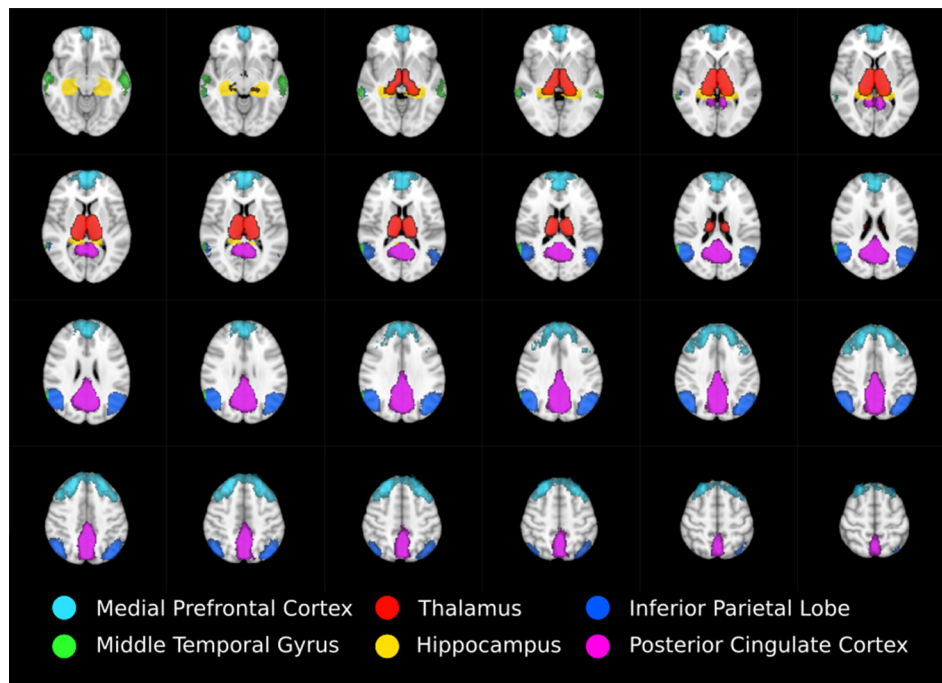


FIGURE 1 | Regions of interest moving from inferior (top left) to superior (bottom right) defining thalamo-DMN white matter tractography masks. DMN components mPFC, MTG, IPL, and PCC were defined using probabilistic template (Wang et al., 2014) while thalamus and hippocampus were defined using the Harvard-Oxford subcortical structural atlas.

thalamus and PCC ($t_{(44)} = -3.91, p < 0.001$), left hippocampus and left IPL ($t_{(44)} = -4.13, p < 0.001$), and left hippocampus and PCC ($t_{(44)} = -3.91, p < 0.001$).

These changes were recapitulated in the MD metric (**Figure 2D**) with significant increases between right hippocampus and PCC ($t_{(44)} = -3.69, p < 0.001$), left thalamus and left hippocampus ($t_{(44)} = -4.34, p < 0.0001$), left thalamus and PCC ($t_{(44)} = -3.63, p = 0.001$) and left hippocampus and PCC ($t_{(44)} = -4.17, p < 0.001$).

Finally, a significant increase in RD (**Figure 2E**) between left hippocampus and PCC ($t_{(44)} = -3.66, p < 0.001$) was also found.

Empirical Measures of Effective and Structural Connectivity Predict Memory Performance

To investigate whether empirical measures of effective and structural connectivity relate to memory, we regressed the diffusivity and GC metrics against the results from a face-name encoding and recognition task using gender, age and motion parameter estimates as covariates of no interest.

The aMCI cohort displayed a significant negative correlation between the integrity of the left thalamo-cortical white matter connectivity and memory in three DMN regions (**Figure 4A**) including IPL ($t_{(24)} = -2.43, p < 0.05$), hippocampus ($t_{(24)} = -2.31, p < 0.05$), and PCC ($t_{(24)} = -2.21, p < 0.05$). The healthy subjects displayed no such relationship.

Conversely, the healthy cohort demonstrated a significant negative correlation between the effective connectivity of IPL and

memory and three other DMN regions (**Figure 4B**) including left MTG ($t_{(24)} = -2.47, p < 0.05$), right IPL ($t_{(24)} = -2.54, p < 0.05$), and PCC ($t_{(24)} = -2.21, p < 0.05$). The same relationship was absent in the aMCI cohort. All results survived multiple-comparison correction with FDR ($q < 0.1$).

DISCUSSION

The appearance of atrophy, tangles and plaques in thalamus is often characterized as a secondary process resulting from atrophy in the hippocampus and the prion-like transmission of pathology along the white matter topography (Raj et al., 2012). But such a view is inconsistent with evidence suggesting that the earliest metabolic changes occur not in hippocampus but in posterior DMN node PCC. Thus, structural deficits in thalamus may be driving early PCC hypometabolism and initiating the cascade of DMN functional anomalies typically associated with early AD. Accordingly, we used a multimodal approach to assess the impact of thalamo-cortico-thalamic feedback loop integrity on DMN functionality in aMCI.

We found significant structural abnormalities in the thalamo-PCC and thalamo-IPL white matter fiber pathways in the aMCI cohort (**Figures 2B–E**). A pronounced reduction in left thalamo-IPL effective connectivity (**Figure 2A**) corresponded with significant thalamo-IPL structural impairment (**Figure 2B**). Critically, the integrity of thalamic white matter and memory was correlated in the aMCI cohort but not in the HCs (**Figure 4A**).

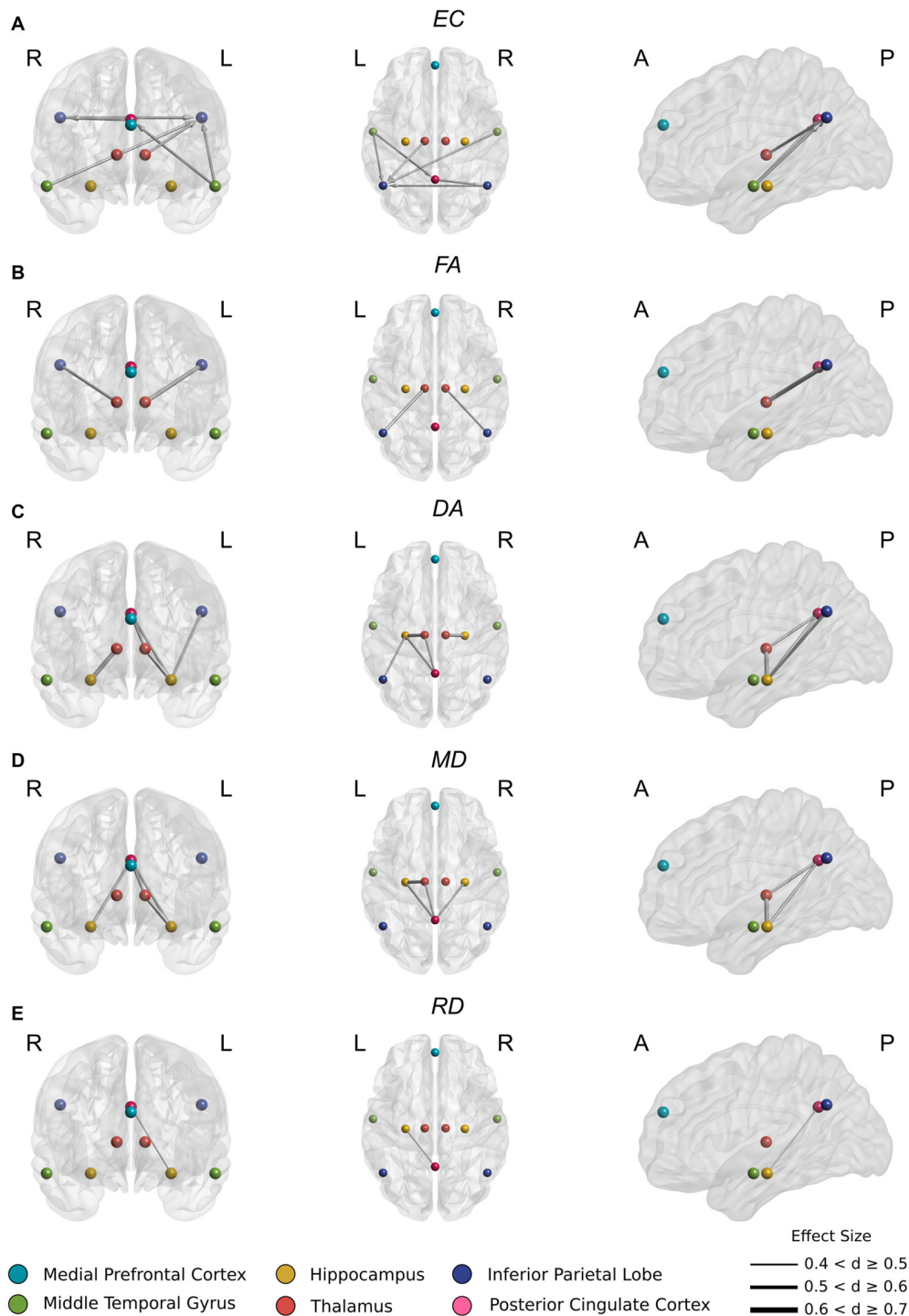


FIGURE 2 | (A) Significantly reduced incoming effective connectivity to left IPL from thalamus and posterior DMN nodes. **(B)** Significantly reduced FA in thalamo-IPL tracts where the magnitude of reduction corresponded to the degree of effective connectivity disruption in **(A)**. **(C)** Significantly reduced DA in the left Papez circuit including hippocampo-thalamus, thalamo-PCC and PCC-hippocampal tracts. **(D)** As in **(C)**, significantly reduced MD in the left Papez circuit. **(E)** Significantly reduced RD in left hippocampo-PCC tracts.

In general, the gradient of structural impairment followed a hippocampo-thalamus-PCC axis consistent with a prion-like dissemination of pathology (Raj et al., 2012) along the major

white matter fiber pathways of the Papez circuit (Papez, 1937). No structural abnormalities were identified between cortical DMN components mPFC, MTG, IPL and PCC, however

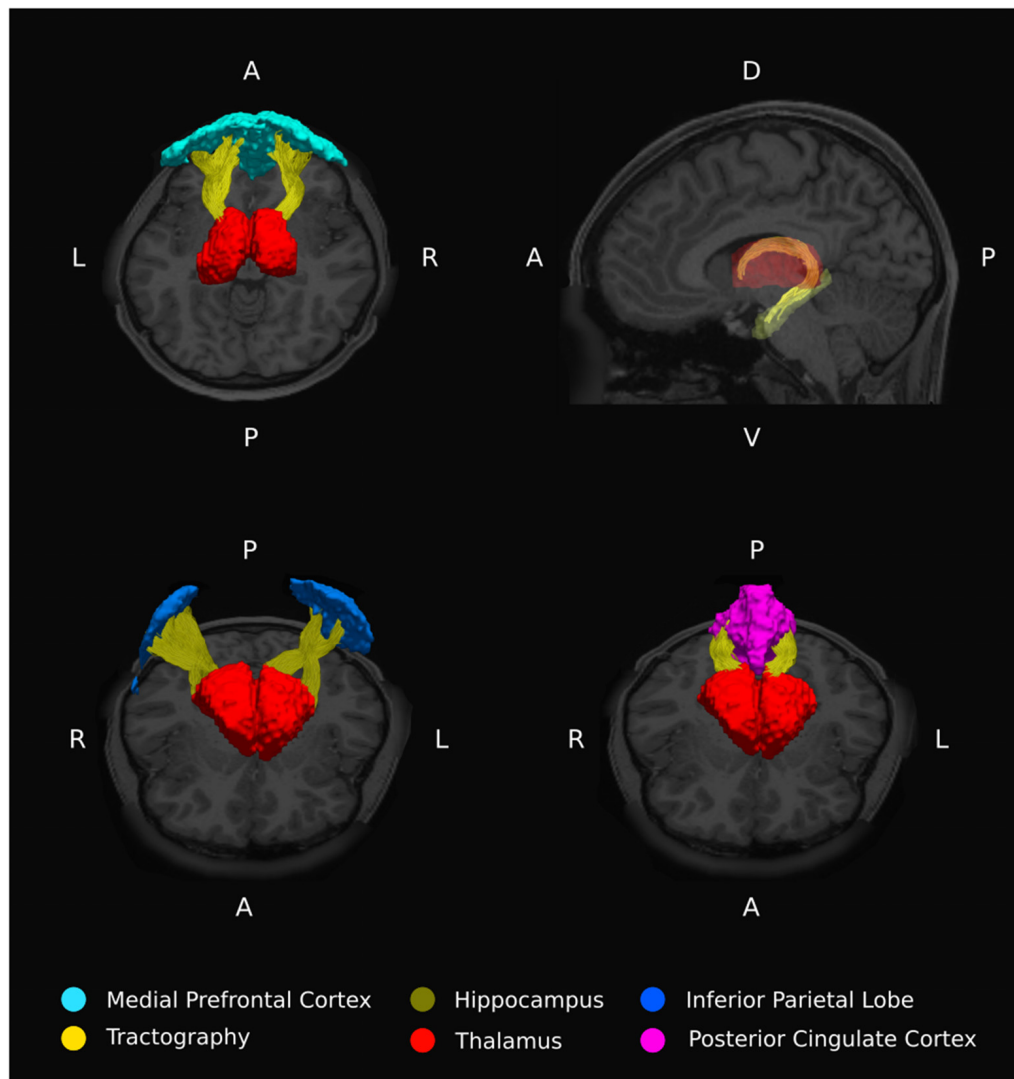


FIGURE 3 | Example thalamo-DMN white matter tracts from a single representative healthy subject.

significant disruption to incoming IPL effective connectivity was observed (**Figure 2A**), and this distinguished HC and aMCI memory performance (**Figure 4B**).

Overall, our findings are broadly suggestive. One interpretation is that disrupted effective connectivity in posterior DMN nodes PCC and IPL is, to some extent, inspired by incipient thalamo-cortical deafferentation. If true, this finding may help explain abnormal task-induced DMN response patterns typically found in aMCI and AD subjects (Pihlajamäki and Sperling, 2009).

Impaired Hippocampo-Thalamo-PCC White Matter Anatomy and Abnormal PCC Effective Connectivity

The current article identified significant structural impairment between fiber pathways connecting hippocampus and thalamus

(**Figures 2C,D**), thalamus and PCC (**Figures 2C,D**), and PCC and hippocampus (**Figures 2C–E**). Measures of left thalamo-cortical structural integrity (including tracts to hippocampus and PCC) correlated with memory performance in the aMCI cohort but not in the HCs (**Figure 4A**).

Impaired structural relations within the hippocampo-thalamo-PCC complex are likely mediated by their close anatomical association. Together these structures comprise a limbic-diencephalic memory network (Nestor et al., 2003) connected through the circuit of Papez (1937). This structure runs from hippocampus through fornix to anterior thalamus via mammillary bodies and onto PCC before returning to hippocampus to complete the circuit. Interestingly, the current study identified a decreasing gradient of structural impairment between hippocampus, thalamus, and PCC, suggesting that structural deafferentation of PCC through

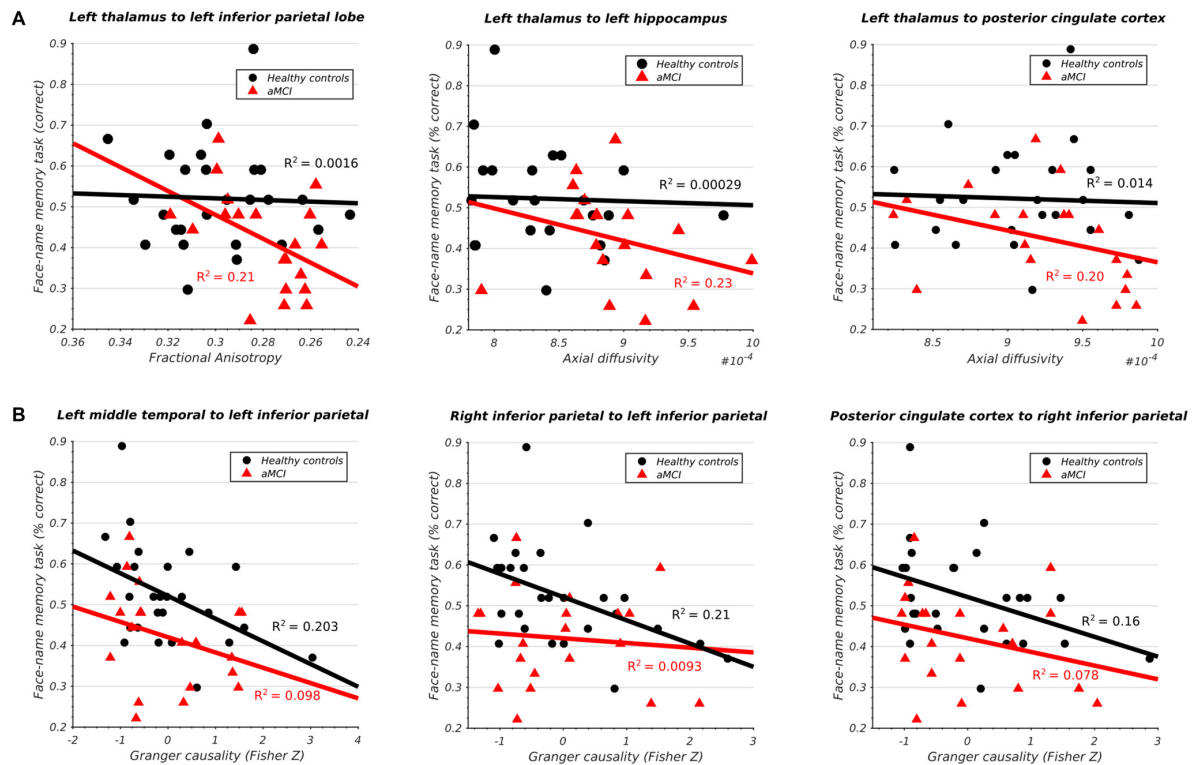


FIGURE 4 | (A) Significant association between structural integrity of thalamo-cortical white matter pathways and memory performance in aMCI subjects. **(B)** Significant association between inferior parietal lobe effective connectivity and memory performance in HCs.

impaired hippocampus and thalamus fiber pathways, likely stems from pathology and atrophy originating in the hippocampal complex. Such a view is consistent with postmortem studies indicating that thalamic nuclei connected to hippocampus are a site of primary degeneration in AD (Xuereb et al., 1991).

Previous work has highlighted a staged disconnection process occurring both along the cingulum bundle between hippocampus and PCC (i.e., the direct route) and within the memory circuit of Papez encompassing thalamic intermediaries (i.e., the indirect route; Villain et al., 2008). Such findings are consistent with early PCC hypometabolism (Matsuda, 2001; Valla et al., 2001; Mosconi et al., 2008; Zhu et al., 2013; Mutlu et al., 2016) where it frequently presents before clinical diagnosis (Minoshima et al., 1997; Johnson et al., 1998) as part of a constellation of metabolic effects focused around medial temporal lobe and thalamus, when memory loss is still a relatively isolated feature (Nestor et al., 2003). Interestingly, PCC hypometabolism appears to correlate with remote hippocampus atrophy early in MCI but transition to both remote and local effects over the course of progression to AD (Teipel and Grothe, 2016).

The current study identified a significant correlation between the structural integrity of hippocampo-thalamus and thalamo-PCC fiber pathways (i.e., the indirect route)

and memory in the aMCI cohort which was absent in the HCs (Figure 4A). A similar pattern was identified in the hippocampo-PCC fiber pathway (i.e., the direct route) however this did not survive correction for multiple comparisons. Dysfunction of structures along the hippocampal output pathways to PCC have been linked to episodic memory impairment (Yakushev et al., 2011).

The PCC's hub status (Hagmann et al., 2008) may predispose to amyloid deposition, atrophy, and hypometabolism (Buckner et al., 2005, 2009) where remote often diffuse damage accumulates as altered PCC connectivity through a form of diaschisis (Meguro et al., 1999; Leech and Sharp, 2014). One suggestion is that direct thalamo-PCC (Figures 2C,D) and distal thalamo-IPL (Figure 2B) white matter structural deficits operate in tandem to initiate a cascade of aberrant effective connectivity in PCC (Figure 2A). Taken together, these findings are consistent with a progressive disconnection of PCC from downstream cortical and subcortical sources with differential effects operating on the direct vs. indirect hippocampo-PCC pathways.

Impaired Thalamo-IPL White Matter Anatomy and Abnormal IPL Effective Connectivity

The current article identified significant impairments in thalamic white matter circuitry serving bilateral IPL where the magnitude of diffusivity change (Figure 2B) correlated with the

intensity and extent of effective connectivity disruption in each hemisphere (**Figure 2A**).

Marked structural deficits were observed in left thalamo-IPL white matter connectivity together with significantly reduced effective connectivity from left thalamus. In the aMCI cohort, measures of reduced thalamo-IPL structural integrity correlated with memory performance (**Figure 4A**). Left thalamo-IPL structural abnormalities were accompanied by widespread decreases in effective connectivity from other DMN regions. Similarly, in right hemisphere, thalamo-IPL structural deficits cooccurred with disrupted incoming and outgoing IPL effective connectivity. Crucially, the relationship between IPL effective connectivity and memory was disrupted in the aMCI subjects but not in the HCs (**Figure 4B**).

Several converging findings implicate the pulvinar nucleus of the thalamus in this dysfunction. The pulvinar nuclei appear to play a role in cortico-cortical communication where they receive driving input from IPL and relay signals back to cortex via ascending thalamo-cortical projections (Saalmann et al., 2012). Since direct cortico-cortical projections far outnumber projections to pulvinar nucleus from cortex, the pulvinar is unlikely to be the primary route for the transfer of cortico-cortical sensory signals, rather, it may act to coordinate interactions between distributed cortical networks as a function of attention (Basso et al., 2005). Interestingly, entorhinal cortex connects directly to pulvinar nucleus via a non-fornical temporopulvinar tract (Saunders et al., 2005; Zarei et al., 2013) which may provide a conduit for the prion-like transsynaptic spread of disease agents originating in hippocampus (Raj et al., 2012). Consistent with this hypothesis, the present study identified significant structural impairment between thalamus and hippocampus (**Figures 2C,D**).

Taken together, these findings are consistent with the idea that disrupted posterior DMN node effective connectivity is, to some extent, mediated by impaired thalamo-cortical white matter circuitry.

Methodological Considerations

Some limitations should be noted. The major weakness of the article is that each thalamic nucleus has specific cortical connections and functions, yet the present analysis uses a *holo-thalamic* approach. It would be more informative to determine whether sub-nuclei show differential causal interactions between specific regions of thalamus and crucial DMN nodes and likewise, whether these connections show varying degrees of structural impairment. Such an approach would reveal the specificity of AD pathology for individual thalamic nuclei. Analyzing the entire thalamus may dilute these results. Our findings should therefore be considered as preliminary evidence warranting further investigation.

It should also be noted that the indirect relationship between fMRI BOLD signal and the underlying neural mechanism is especially problematic when applying GC and should be noted as a weakness in the present study. First, the study's sampling rate (repetition time or TR) of 2 s is considerably slower than the millisecond temporal resolution of the neuronal activity we seek to qualify. Second, the temporal precedence assumptions

of GC can be violated by regional differences in the latency of the hemodynamic response (Handwerker et al., 2004; Friston, 2011). Since neurovascular coupling can be altered in complex ways by disease, the likelihood of such an event is magnified in the aMCI patient cohort (Handwerker et al., 2012). One typical scenario, is that region X causally influences Y at the neuronal level but has a longer time-to-peak in its HRF due to pathology of the neurovasculature. Thus, GC analysis of BOLD signal may incorrectly suggest that Y is causally implicated in causing X. Simulations show that GC performs well when the HRF delay between regions is short (Deshpande et al., 2010; Schippers et al., 2011) however sufficiently fast sampling, on par with the neuronal delays themselves, is required to ensure GC is fully invariant to HRF latency (Seth et al., 2013). Other simulations suggest that the relationship between GC at the neuronal level and GC at the fMRI level is reasonably preserved over a range of sampling rates and convolution parameters (Wen et al., 2013). Whatever the case, sub-second temporal resolutions have been made available (Feinberg et al., 2010; Feinberg and Yacoub, 2012) and are standard as part of the Human Connectome Project (Van Essen et al., 2013). The most recent advances enable a temporal resolution as fast as 50 ms (Boyacioglu and Barth, 2013).

Critically, GC makes no claims regarding the underlying physical mechanisms responsible for the observed differences in causal relationships between regions. In contrast, the dynamic causal modeling approach (DCM; Di and Biswal, 2014) explicitly specifies dynamic effective relationships at the neuronal level, allowing the most likely structural model for generating the observed data to be identified. Applying DCM in future studies will help clarify thalamic involvement in posterior DMN dysfunction.

The choice of CSD-based tractography reflects the growing recognition that assumptions underlying the DTI model may not always be met in practice (Wheeler-Kingshott and Cercignani, 2009; Jones, 2010; Jones and Cercignani, 2010). The DTI model can only capture a unitary fiber direction within a single voxel despite observations that 90% of the brain is composed of multiple crossing fibers (Jeurissen et al., 2013). For this reason CSD attempts to map several fiber directions per voxel by taking advantage of the high number of diffusion encoding directions and large b-values acquired using the HARDI acquisition protocol (Tournier et al., 2007, 2008; Mielke et al., 2012; Farquharson et al., 2013). Using large b-values has an additional advantage. By allowing a sufficiently long diffusion path to be measured water molecules are more likely to collide with their container. This may be relevant in patients with neurodegenerative disorders who have increased permeability of membranes, greater extracellular space due to axonal atrophy, demyelination and glial pathology (Acosta-Cabronero and Nestor, 2014). To date, only a handful of tractography studies have utilized HARDI data and large b-values (Thiebaut de Schotten et al., 2011; Meng et al., 2013; Yeatman et al., 2014; Xie et al., 2015) and only one specifically in clinical aMCI and AD (Kehoe et al., 2015).

The absence of indirect biomarkers of AD pathology (CSF biomarkers and/or amyloid PET imaging) should also be acknowledged as a weakness in the present article.

CONCLUSION

The dynamic nature of thalamo-cortical dialog suggests that abnormalities in DMN operation may best be understood from the perspective of thalamic dysfunction. The present study employed diffusion imaging and effective connectivity to clarify the relationship between the physical integrity of thalamic white matter projections and the activity of the DMN. Significant changes in the diffusivity metrics of thalamic white matter projection tracts to hippocampus, PCC and IPL (**Figures 2B–D**) were identified. Effective connectivity changes corresponding to the same regions were also observed (**Figure 2A**). Interestingly, no structural deficits were found between DMN nodes suggesting that early changes in DMN activity could be a result of impaired thalamo-cortical structural integrity.

Such a conclusion is supported by previous resting state MEG (Garcés et al., 2014), EEG (Schreckenberger et al., 2004; Garcés et al., 2013; Moretti, 2015) and fMRI (Greicius et al., 2004; Sorg et al., 2007; Damoiseaux et al., 2012) studies citing disruption in posterior thalamo-cortical alpha sources. Significant evidence suggests that thalamo-cortical circuitry underlie the generation and modulation of alpha and theta rhythms and that average power is attenuated in these frequency bands for MCI and AD subjects (Jeong, 2004; Koenig et al., 2005; Jelles et al., 2008; Park et al., 2008). Several recent modeling studies have proposed a candidate mechanism citing impairment to thalamic reticular fibers in MCI and AD as the source of the dysfunction (Bhattacharya et al., 2011, 2013; Li et al., 2011; Abuhassan et al., 2014).

A corollary of this discussion is the extent to which cortical activity is dependent on cortico-cortical versus thalamo-cortical connections. It has been suggested that thalamic nuclei coordinate distributed cortical regions through cortico-thalamo-cortical pathways. Abnormalities originating in thalamic to PCC and IPL white matter may therefore be sufficient to engender posterior DMN dysfunction without appealing to comparable deficits in cortico-cortical tracts between DMN nodes. Such a view is consistent with the anatomy and timeline of pathogenesis with thalamic nuclei demonstrating pathology an earlier stage of the disease than cortex. Cortical atrophy

may therefore be in response to thalamic white matter disruption with commensurate causal abnormalities occurring in response to changes in thalamo-cortical signaling rather than being instigated by structural changes within the cortex. Importantly, the present study is unable to confirm this hypothesis. Other scenarios, in which cortical pathology is causing a degeneration of thalamo-cortical tracts is also possible or likewise, a parallel disruption in both thalamus and cortex.

Overall, these results provide a compelling and previously unexplored physical basis for posterior DMN dysfunction and abnormal fMRI task-induced deactivation response patterns in aMCI and AD patients and underscore the need to consider neurodegenerative changes within a wider system context including contributions of both cortical and subcortical thalamic components. This work complements a growing body of evidence that suggests effective connectivity is disrupted in neurodegenerative disorders such as aMCI and AD and that these changes are underpinned by structural deficits. For these reasons, joint effective and structural studies will play an increasingly important role in the future as we seek to understand how pathological changes in structural connectivity are reflected in altered network effective connectivity.

AUTHOR CONTRIBUTIONS

TA: analysis, manuscript preparation. DC, ALWB and LM: supervisory support. EK, DF, BL, RAK and DL: data collection.

ACKNOWLEDGMENTS

The research was funded by a Department for Employment and Learning Northern Ireland PhD studentship and the data collection was supported by the Northern Ireland Department for Education and Learning under the Strengthening the All Island Research Base programme. ALWB was funded in part by the Science Foundation Ireland Stokes Programme (07/SK/B1214a). RAK and BL funded in part by the Health Research Board (Ireland).

REFERENCES

- Abuhassan, K., Coyle, D., and Maguire, L. (2014). Compensating for thalamocortical synaptic loss in Alzheimer's disease. *Front. Comput. Neurosci.* 8:65. doi: 10.3389/fncom.2014.00065
- Acosta-Cabrero, J., and Nestor, P. J. (2014). Diffusion tensor imaging in Alzheimer's disease: insights into the limbic-diencephalic network and methodological considerations. *Front. Aging Neurosci.* 6:266. doi: 10.3389/fnagi.2014.00266
- Aggleton, J. P., O'Mara, S. M., Vann, S. D., Wright, N. F., Tsanov, M., Erichsen, J. T., et al. (2010). Hippocampal-anterior thalamic pathways for memory: uncovering a network of direct and indirect actions. *Eur. J. Neurosci.* 31, 2292–2307. doi: 10.1111/j.1460-9568.2010.07251.x
- Andersson, J. L. R., Jenkinson, M., and Smith, S. (2007). Non-linear registration aka Spatial normalisation FMRIB Technical Report TR07JA2. In practice. Available online at: <https://www.fmrib.ox.ac.uk/datasets/techrep/tr07ja2/tr07ja2.pdf>
- Andersson, J. L. R., and Sotiropoulos, S. N. (2016). An integrated approach to correction for off-resonance effects and subject movement in diffusion MR imaging. *Neuroimage* 125, 1063–1078. doi: 10.1016/j.neuroimage.2015.10.019
- Basso, M. A., Uhlrich, D., and Bickford, M. E. (2005). Cortical function: a view from the thalamus. *Neuron* 45, 485–488. doi: 10.1016/j.neuron.2005.01.035
- Beckmann, C. F., and Smith, S. M. (2004). Probabilistic independent component analysis for functional magnetic resonance imaging. *IEEE Trans. Med. Imaging* 23, 137–152. doi: 10.1109/tmi.2003.822821
- Bhattacharya, B. S., Cakir, Y., Serap-Sengor, N., Maguire, L., and Coyle, D. (2013). Model-based bifurcation and power spectral analyses of thalamocortical alpha rhythm slowing in Alzheimer's disease. *Neurocomputing* 115, 11–22. doi: 10.1016/j.neucom.2012.10.023
- Bhattacharya, B. S., Coyle, D., and Maguire, L. P. (2011). A thalamo-cortico-thalamic neural mass model to study alpha rhythms in Alzheimer's disease. *Neural Netw.* 24, 631–645. doi: 10.1016/j.neunet.2011.02.009

- Boyacioglu, R., and Barth, M. (2013). Generalized inverse imaging (GIN): ultrafast fMRI with physiological noise correction. *Magn. Reson. Med.* 70, 962–971. doi: 10.1002/mrm.24528
- Braak, H., and Braak, E. (1991a). Alzheimer's disease affects limbic nuclei of the thalamus. *Acta Neuropathol.* 81, 261–268. doi: 10.1007/BF00305867
- Braak, H., and Braak, E. (1991b). Neuropathological staging of Alzheimer-related changes. *Acta Neuropathol.* 82, 239–259. doi: 10.1007/bf00308809
- Braak, H., and Braak, E. (1995). Staging of Alzheimer's disease-related neurofibrillary changes. *Neurobiol. Aging* 16, 271–278. doi: 10.1016/0197-4580(95)00021-6
- Buckner, R. L., Andrews-Hanna, J. R., and Schacter, D. L. (2008). The brain's default network: anatomy, function, and relevance to disease. *Ann. N Y Acad. Sci.* 1124, 1–38. doi: 10.1196/annals.1440.011
- Buckner, R. L., Sepulcre, J., Talukdar, T., Krienen, F. M., Liu, H., Hedden, T., et al. (2009). Cortical hubs revealed by intrinsic functional connectivity: mapping, assessment of stability, and relation to Alzheimer's disease. *J. Neurosci.* 29, 1860–1873. doi: 10.1523/JNEUROSCI.5062-08.2009
- Buckner, R. L., Snyder, A. Z., Shannon, B. J., LaRossa, G., Sachs, R., Fotenos, A. F., et al. (2005). Molecular, structural, and functional characterization of Alzheimer's disease: evidence for a relationship between default activity, amyloid, and memory. *J. Neurosci.* 25, 7709–7717. doi: 10.1523/JNEUROSCI.2177-05.2005
- Burns, A., and Iliffe, S. (2009). Alzheimer's disease. *BMJ* 338:b158. doi: 10.1136/bmj.b158
- Cantero, J. L., Atienza, M., Gomez-Herrero, G., Cruz-Vadell, A., Gil-Neciga, E., Rodriguez-Romero, R., et al. (2009). Functional integrity of thalamocortical circuits differentiates normal aging from mild cognitive impairment. *Hum. Brain Mapp.* 30, 3944–3957. doi: 10.1002/hbm.20819
- Cherubini, A., Pèran, P., Spoletini, I., Di Paola, M., Di Iulio, F., Hagberg, G. E., et al. (2010). Combined volumetry and DTI in subcortical structures of mild cognitive impairment and Alzheimer's disease patients. *J. Alzheimers Dis.* 19, 1273–1282. doi: 10.3233/JAD-2010-091186
- Chételat, G., Landeau, B., Eustache, F., Mézenge, F., Viader, F., de la Sayette, V., et al. (2005). Using voxel-based morphometry to map the structural changes associated with rapid conversion in MCI: a longitudinal MRI study. *Neuroimage* 27, 934–946. doi: 10.1016/j.neuroimage.2005.05.015
- Chhatwal, J. P., Schultz, A. P., Johnson, K., Benzinger, T. L. S., Jack, C. Jr., Ances, B. M., et al. (2013). Impaired default network functional connectivity in autosomal dominant Alzheimer disease. *Neurology* 81, 736–744. doi: 10.1212/WNL.0b013e3182a1aaf6
- Chiang, G. C., Zhan, W., Schuff, N., and Weiner, M. W. (2012). White matter alterations in cognitively normal apoEε2 carriers: insight into Alzheimer resistance? *Am. J. Neuroradiol.* 33, 1392–1397. doi: 10.3174/ajnr.a2984
- Damoiseaux, J. S., Prater, K. E., Miller, B. L., and Greicius, M. D. (2012). Functional connectivity tracks clinical deterioration in Alzheimer's disease. *Neurobiol. Aging* 33, 828.e19–828.30. doi: 10.1016/j.neurobiolaging.2011.06.024
- Damoiseaux, J. S., Rombouts, S. A. R. B., Barkhof, F., Scheltens, P., Stam, C. J., Smith, S. M., et al. (2006). Consistent resting-state networks across healthy subjects. *Proc. Natl. Acad. Sci. U S A* 103, 13848–13853. doi: 10.1073/pnas.0601417103
- de Jong, L. W., van der Hiele, K., Veer, I. M., Houwing, J. J., Westendorp, R. G. J., Bollen, E. L. E. M., et al. (2008). Strongly reduced volumes of putamen and thalamus in Alzheimer's disease: an MRI study. *Brain* 131, 3277–3285. doi: 10.1093/brain/awn278
- Deshpande, G., Sathian, K., and Hu, X. (2010). Effect of hemodynamic variability on Granger causality analysis of fMRI. *Neuroimage* 52, 884–896. doi: 10.1016/j.neuroimage.2009.11.060
- Di, X., and Biswal, B. B. (2014). Modulatory interactions between the default mode network and task positive networks in resting-state. *PeerJ* 2:e367. doi: 10.7717/peerj.367
- Drzezga, A., Becker, J. A., Van Dijk, K. R., Sreenivasan, A., Talukdar, T., Sullivan, C., et al. (2011). Neuronal dysfunction and disconnection of cortical hubs in non-demented subjects with elevated amyloid burden. *Brain* 134, 1635–1646. doi: 10.1093/brain/awr066
- Erwin, R. J., Gur, R. C., Gur, R. E., Skolnick, B., Mawhinney-Hee, M., and Smailis, J. (1992). Facial emotion discrimination: I. Task construction and behavioral findings in normal subjects. *Psychiatry Res.* 42, 231–240. doi: 10.1016/0165-1781(92)90115-j
- Eysenck, H. J., and Eysenck, S. B. G. (1994). *Manual for the Eysenck Personality Questionnaire (EPQ-R Adult)*. San Diego, CA: EdITS/Educational and Industrial Testing Service.
- Farquharson, S., Tournier, J.-D., Calamante, F., Fabbini, G., Schneider-Kolsky, M., Jackson, G. D., et al. (2013). White matter fiber tractography: why we need to move beyond DTI. *J. Neurosurg.* 118, 1367–1377. doi: 10.3171/2013.2.JNS121294
- Feinberg, D. A., Moeller, S., Smith, S. M., Auerbach, E., Ramanna, S., Glasser, M. F., et al. (2010). Multiplexed echo planar imaging for sub-second whole brain fmri and fast diffusion imaging. *PLoS One* 5:e15710. doi: 10.1371/journal.pone.0015710
- Feinberg, D. A., and Yacoub, E. (2012). The rapid development of high speed, resolution and precision in fMRI. *Neuroimage* 62, 720–725. doi: 10.1016/j.neuroimage.2012.01.049
- Fernández-Espejo, D., Soddu, A., Cruse, D., Palacios, E. M., Junque, C., Vanhaudenhuyse, A., et al. (2012). A role for the default mode network in the bases of disorders of consciousness. *Ann. Neurol.* 72, 335–343. doi: 10.1002/ana.23635
- Ferrarini, L., Palm, W. M., Olofsen, H., van der Landen, R., Jan Blauw, G., Westendorp, R. G. J., et al. (2008). MMSE scores correlate with local ventricular enlargement in the spectrum from cognitively normal to Alzheimer disease. *Neuroimage* 39, 1832–1838. doi: 10.1016/j.neuroimage.2007.11.003
- Fletcher, E., Carmichael, O., Pasternak, O., Maier-Hein, K. H., and DeCarli, C. (2014). Early brain loss in circuits affected by Alzheimer's disease is predicted by fornix microstructure but may be independent of gray matter. *Front. Aging Neurosci.* 6:106. doi: 10.3389/fnagi.2014.00106
- Folstein, M. F., Folstein, S. E., and McHugh, P. R. (1975). “Mini-mental state”. A practical method for grading the cognitive state of patients for the clinician. *J. Psychiatr. Res.* 12, 189–198. doi: 10.1016/0022-3956(75)90026-6
- Friston, K. J. (2011). Functional and effective connectivity: a review. *Brain Connect.* 1, 13–36. doi: 10.1089/brain.2011.0008
- Garcés, P., Angel Pineda-Pardo, J., Canuet, L., Aurteneixe, S., López, M. E., Marcos, A., et al. (2014). The Default Mode Network is functionally and structurally disrupted in amnesic mild cognitive impairment—a bimodal MEG-DTI study. *Neuroimage Clin.* 6, 214–221. doi: 10.1016/j.nicl.2014.09.004
- Garcés, P., Vicente, R., Wibrál, M., Pineda-Pardo, J. Á., López, M. E., Aurteneixe, S., et al. (2013). Brain-wide slowing of spontaneous α rhythms in mild cognitive impairment. *Front. Aging Neurosci.* 5:100. doi: 10.3389/fnagi.2013.00100
- Granger, C. W. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica* 37, 424–438. doi: 10.2307/1912791
- Greicius, M. D., Krasnow, B., Reiss, A. L., and Menon, V. (2003). Functional connectivity in the resting brain: a network analysis of the default mode hypothesis. *Proc. Natl. Acad. Sci. U S A* 100, 253–258. doi: 10.1073/pnas.0135058100
- Greicius, M. D., Srivastava, G., Reiss, A. L., and Menon, V. (2004). Default-mode network activity distinguishes Alzheimer's disease from healthy aging: evidence from functional MRI. *Proc. Natl. Acad. Sci. U S A* 101, 4637–4642. doi: 10.1073/pnas.0308627101
- Hagmann, P., Cammoun, L., Gigandet, X., Meuli, R., Honey, C. J., Wedeen, V. J., et al. (2008). Mapping the structural core of human cerebral cortex. *PLoS Biol.* 6:e159. doi: 10.1371/journal.pbio.0060159
- Handwerker, D. A., Gonzalez-Castillo, J., D'Esposito, M., and Bandettini, P. A. (2012). The continuing challenge of understanding and modeling hemodynamic variation in fMRI. *Neuroimage* 62, 1017–1023. doi: 10.1016/j.neuroimage.2012.02.015
- Handwerker, D. A., Ollinger, J. M., and D'Esposito, M. (2004). Variation of BOLD hemodynamic responses across subjects and brain regions and their effects on statistical analyses. *Neuroimage* 21, 1639–1651. doi: 10.1016/j.neuroimage.2003.11.029
- Hedden, T., Van Dijk, K. R. A., Becker, J. A., Mehta, A., Sperling, R. A., Johnson, K. A., et al. (2009). Disruption of functional connectivity in clinically normal older adults harboring amyloid burden. *J. Neurosci.* 29, 12686–12694. doi: 10.1523/JNEUROSCI.3189-09.2009
- Huppert, F. A., Brayne, C., Gill, C., Paykel, E. S., and Beardsall, L. (1995). CAMCOG—a concise neuropsychological test to assist dementia diagnosis: socio-demographic determinants in an elderly population sample. *Br. J. Clin. Psychol.* 34, 529–541. doi: 10.1111/j.2044-8260.1995.tb01487.x

- Hyvärinen, A. (1999). Fast and robust fixed-point algorithms for independent component analysis. *IEEE Trans. Neural Netw.* 10, 626–634. doi: 10.1109/72.761722
- Jelles, B., Scheltens, P., van der Flier, W. M., Jonkman, E. J., da Silva, F. H. L., and Stam, C. J. (2008). Global dynamical analysis of the EEG in Alzheimer's disease: frequency-specific changes of functional interactions. *Clin. Neurophysiol.* 119, 837–841. doi: 10.1016/j.clinph.2007.12.002
- Jenkinson, M., Bannister, P., Brady, M., and Smith, S. (2002). Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage* 17, 825–841. doi: 10.1016/s1053-8119(02)91132-8
- Jeong, J. (2004). EEG dynamics in patients with Alzheimer's disease. *Clin. Neurophysiol.* 115, 1490–1505. doi: 10.1016/j.clinph.2004.01.001
- Jeurissen, B., Leemans, A., Tournier, J.-D., Jones, D. K., and Sijbers, J. (2013). Investigating the prevalence of complex fiber configurations in white matter tissue with diffusion magnetic resonance imaging. *Hum. Brain Mapp.* 34, 2747–2766. doi: 10.1002/hbm.22099
- Johnson, K. A., Jones, K., Holman, B. L., Becker, J. A., Spiers, P. A., Satlin, A., et al. (1998). Preclinical prediction of Alzheimer's disease using SPECT. *Neurology* 50, 1563–1571. doi: 10.1212/WNL.50.6.1563
- Jones, D. K. (2010). Challenges and limitations of quantifying brain connectivity *in vivo* with diffusion MRI. *Imaging Med.* 2, 341–355. doi: 10.2217/iim.10.21
- Jones, D. K., and Cercignani, M. (2010). Twenty-five pitfalls in the analysis of diffusion MRI data. *NMR Biomed.* 23, 803–820. doi: 10.1002/nbm.1543
- Jones, D. T., Mateen, F. J., Lucchinetti, C. F., Jack, C. R. Jr., and Welker, K. M. (2011). Default mode network disruption secondary to a lesion in the anterior thalamus. *Arch. Neurol.* 68, 242–247. doi: 10.1001/archneurol.2010.259
- Kehoe, E. G., Farrell, D., Metzler-Baddeley, C., Lawlor, B. A., Kenny, R. A., Lyons, D., et al. (2015). Fornix white matter is correlated with resting-state functional connectivity of the thalamus and hippocampus in healthy aging but not in mild cognitive impairment—a preliminary study. *Front. Aging Neurosci.* 7:10. doi: 10.3389/fnagi.2015.00010
- Knight, W. D., Kim, L. G., Douiri, A., Frost, C., Rossor, M. N., and Fox, N. C. (2011a). Acceleration of cortical thinning in familial Alzheimer's disease. *Neurobiol. Aging* 32, 1765–1773. doi: 10.1016/j.neurobiolaging.2009.11.013
- Knight, W. D., Okello, A. A., Ryan, N. S., Turkheimer, F. E., Rodríguez Martínez De Llano, S., Edison, P., et al. (2011b). Carbon-11-Pittsburgh compound B positron emission tomography imaging of amyloid deposition in presenilin 1 mutation carriers. *Brain* 134, 293–300. doi: 10.1093/brain/awq310
- Koenig, T., Prichep, L., Dierks, T., Hubl, D., Wahlund, L. O., John, E. R., et al. (2005). Decreased EEG synchronization in Alzheimer's disease and mild cognitive impairment. *Neurobiol. Aging* 26, 165–171. doi: 10.1016/j.neurobiolaging.2004.03.008
- Leech, R., and Sharp, D. J. (2014). The role of the posterior cingulate cortex in cognition and disease. *Brain* 137, 12–32. doi: 10.1093/brain/awt162
- Li, X., Coyle, D., Maguire, L., Watson, D. R., and McGinnity, T. M. (2011). Gray matter concentration and effective connectivity changes in Alzheimer's disease: a longitudinal structural MRI study. *Neuroradiology* 53, 733–748. doi: 10.1007/s00234-010-0795-1
- Luo, Q., Lu, W., Cheng, W., Valdes-Sosa, P. A., Wen, X., Ding, M., et al. (2013). Spatio-temporal Granger causality: a new framework. *Neuroimage* 79, 241–263. doi: 10.1016/j.neuroimage.2013.04.091
- Matsuda, H. (2001). Cerebral blood flow and metabolic abnormalities in Alzheimer's disease. *Ann. Nucl. Med.* 15, 85–92. doi: 10.1007/BF02988596
- Meguro, K., Blaizot, X., Kondoh, Y., Le Mestric, C., Baron, J. C., and Chavoix, C. (1999). Neocortical and hippocampal glucose hypometabolism following neurotoxic lesions of the entorhinal and perirhinal cortices in the non-human primate as shown by PET. Implications for Alzheimer's disease. *Brain* 122, 1519–1531. doi: 10.1093/brain/122.8.1519
- Meng, X., Jun, C., Wang, Q., Zhang, X., Li, Z., Li, Q., et al. (2013). High b-value diffusion tensor imaging of the remote white matter and white matter of obstructive unilateral cerebral arterial regions. *Clin. Radiol.* 68, 815–822. doi: 10.1016/j.crad.2013.03.009
- Metzler-Baddeley, C., Hunt, S., Jones, D. K., Leemans, A., Aggleton, J. P., and O'Sullivan, M. J. (2012a). Temporal association tracts and the breakdown of episodic memory in mild cognitive impairment. *Neurology* 79, 2233–2240. doi: 10.1212/WNL.0b013e31827689e8
- Metzler-Baddeley, C., Jones, D. K., Steventon, J., Westacott, L., Aggleton, J. P., and O'Sullivan, M. (2012b). Cingulum microstructure predicts cognitive control in older age and mild cognitive impairment. *J. Neurosci.* 32, 17612–17619. doi: 10.1523/JNEUROSCI.3299-12.2012
- Mielke, M. M., Okonkwo, O. C., Oishi, K., Mori, S., Tighe, S., Miller, M. L., et al. (2012). Fornix integrity and hippocampal volume predict memory decline and progression to Alzheimer's disease. *Alzheimers Dement.* 8, 105–113. doi: 10.1016/j.jalz.2011.05.2416
- Minka, T. P. (2001). "Automatic choice of dimensionality for PCA," in *Advances in Neural Information Processing Systems*, eds S. Becker, S. Thrun and K. Obermayer (Cambridge, MA: MIT Press), 598–604.
- Minoshima, S., Foster, N. L., and Kuhl, D. E. (1994). Posterior cingulate cortex in Alzheimer's disease. *Lancet* 344:895. doi: 10.1007/978-1-4899-6704-6_17
- Minoshima, S., Giordani, B., Berent, S., Frey, K. A., Foster, N. L., and Kuhl, D. E. (1997). Metabolic reduction in the posterior cingulate cortex in very early Alzheimer's disease. *Ann. Neurol.* 42, 85–94. doi: 10.1002/ana.410420114
- Mintun, M. A., Larossa, G. N., Sheline, Y. I., Dence, C. S., Lee, S. Y., Mach, R. H., et al. (2006). [¹¹C]PIB in a nondemented population: potential antecedent marker of Alzheimer disease. *Neurology* 67, 446–452. doi: 10.1212/01.WNL.0000228230.26044.a4
- Moretti, D. V. (2015). Theta and α EEG frequency interplay in subjects with mild cognitive impairment: evidence from EEG, MRI, and SPECT brain modifications. *Front. Aging Neurosci.* 7:31. doi: 10.3389/fnagi.2015.00031
- Mormino, E. C., Smiljic, A., Hayenga, A. O., Onami, S. H., Greicius, M. D., Rabinovici, G. D., et al. (2011). Relationships between β -amyloid and functional connectivity in different components of the default mode network in aging. *Cereb. Cortex* 21, 2399–2407. doi: 10.1093/cercor/bhr025
- Morris, J. C., Heyman, A., Mohs, R. C., Hughes, J. P., van Belle, G., Fillenbaum, G., et al. (1989). The consortium to establish a registry for Alzheimer's disease (CERAD). Part I. Clinical and neuropsychological assessment of Alzheimer's disease. *Neurology* 39, 1159–1165. doi: 10.1212/WNL.39.9.1159
- Mosconi, L., Tsui, W. H., Herholz, K., Pupi, A., Drzezga, A., Lucignani, G., et al. (2008). Multicenter standardized 18F-FDG PET diagnosis of mild cognitive impairment, Alzheimer's disease, and other dementias. *J. Nucl. Med.* 49, 390–398. doi: 10.2967/jnumed.107.045385
- Mutlu, J., Landeau, B., Tomadesso, C., de Flores, R., Mézenge, F., de La Sayette, V., et al. (2016). Connectivity disruption, atrophy and hypometabolism within posterior cingulate networks in Alzheimer's disease. *Front. Neurosci.* 10:582. doi: 10.3389/fnins.2016.00582
- Nestor, P. J., Fryer, T. D., Smielewski, P., and Hodges, J. R. (2003). Limbic hypometabolism in Alzheimer's disease and mild cognitive impairment. *Ann. Neurol.* 54, 343–351. doi: 10.1002/ana.10669
- Papez, J. (1937). A proposed mechanism of emotion. *Arch. Neurol. Psychiatry* 38, 725–743. doi: 10.1001/archneurpsyc.1937.02260220069003
- Park, Y.-M., Che, H.-J., Im, C.-H., Jung, H.-T., Bae, S.-M., and Lee, S.-H. (2008). Decreased EEG synchronization and its correlation with symptom severity in Alzheimer's disease. *Neurosci. Res.* 62, 112–127. doi: 10.1016/j.neures.2008.06.009
- Pasternak, O., Sochen, N., Gur, Y., Intrator, N., and Assaf, Y. (2009). Free water elimination and mapping from diffusion MRI. *Magn. Reson. Med.* 62, 717–730. doi: 10.1002/mrm.22055
- Patel, K. T., Stevens, M. C., Pearson, G. D., Winkler, A. M., Hawkins, K. A., Skudlarski, P., et al. (2013). Default mode network activity and white matter integrity in healthy middle-aged ApoE4 carriers. *Brain Imaging Behav.* 7, 60–67. doi: 10.1007/s11682-012-9187-y
- Pedro, T., Weiler, M., Yasuda, C. L., D'Abreu, A., Damasceno, B. P., Cendes, F., et al. (2012). Volumetric brain changes in thalamus, corpus callosum and medial temporal structures: mild alzheimer's disease compared with amnesic mild cognitive impairment. *Dement. Geriatr. Cogn. Disord.* 34, 149–155. doi: 10.1159/000342118
- Petersen, R. C. (2004). Mild cognitive impairment as a diagnostic entity. *J. Intern. Med.* 256, 183–194. doi: 10.1111/j.1365-2796.2004.01388.x
- Petersen, R. C., Smith, G. E., Waring, S. C., Ivnik, R. J., Tangalos, E. G., and Kokmen, E. (1999). Mild cognitive impairment: clinical characterization and outcome. *Arch. Neurol.* 56, 303–308. doi: 10.1001/archneur.56.3.303

- Pihlajamäki, M., and Sperling, R. A. (2009). Functional MRI assessment of task-induced deactivation of the default mode network in Alzheimer's disease and at-risk older individuals. *Behav. Neurol.* 21, 77–91. doi: 10.3233/BEN-2009-0231
- Prussmann, K. P., Weiger, M., Scheidegger, M. B., and Boesiger, P. (1999). SENSE: sensitivity encoding for fast MRI. *Magn. Reson. Med.* 42, 952–962. doi: 10.1002/(sici)1522-2594(199911)42:5<952::aid-mrm16>3.3.co;2-j
- Raichle, M. E., MacLeod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A., and Shulman, G. L. (2001). A default mode of brain function. *Proc. Natl. Acad. Sci. U S A* 98, 676–682. doi: 10.1073/pnas.98.2.676
- Raj, A., Kuceyeski, A., and Weiner, M. (2012). A network diffusion model of disease progression in dementia. *Neuron* 73, 1204–1215. doi: 10.1016/j.neuron.2011.12.040
- Rami, L., Valls-Pedret, C., Bartrés-Faz, D., Caprile, C., Solé-Padullés, C., Castellví, M., et al. (2011). Cognitive reserve questionnaire. Scores obtained in a healthy elderly population and in one with Alzheimer's disease. *Rev. Neurol.* 52, 195–201.
- Roh, J. H., Qiu, A., Seo, S. W., Soon, H. W., Kim, J. H., Kim, G. H., et al. (2011). Volume reduction in subcortical regions according to severity of Alzheimer's disease. *J. Neurol.* 258, 1013–1020. doi: 10.1007/s00415-010-5872-1
- Roses, A. D. (1996). Apolipoprotein E alleles as risk factors in Alzheimer's disease. *Annu. Rev. Med.* 47, 387–400. doi: 10.1146/annurev.med.47.1.387
- Ryan, N. S., Keihaninejad, S., Shakespeare, T. J., Lehmann, M., Crutch, S. J., Malone, I. B., et al. (2013). Magnetic resonance imaging evidence for presymptomatic change in thalamus and caudate in familial Alzheimer's disease. *Brain* 136, 1399–1414. doi: 10.1093/brain/awt065
- Saalmann, Y. B., Pinsk, M. A., Wang, L., Li, X., and Kastner, S. (2012). The pulvinar regulates information transmission between cortical areas based on attention demands. *Science* 337, 753–756. doi: 10.1126/science.1223082
- Saunders, R. C., Mishkin, M., and Aggleton, J. P. (2005). Projections from the entorhinal cortex, perirhinal cortex, presubiculum and parasubiculum to the medial thalamus in macaque monkeys: identifying different pathways using disconnection techniques. *Exp. Brain Res.* 167, 1–16. doi: 10.1007/s00221-005-2361-3
- Schelter, B., Winterhalder, M., and Timmer, J. (2006). *Handbook of Time Series Analysis*. Berlin: Wiley-VCH.
- Schippers, M. B., Renken, R., and Keysers, C. (2011). The effect of intra- and inter-subject variability of hemodynamic responses on group level Granger causality analyses. *Neuroimage* 57, 22–36. doi: 10.1016/j.neuroimage.2011.02.008
- Schreckenberger, M., Lange-Asschenfeldt, C., Lochmann, M., Mann, K., Siessmeier, T., Buchholz, H. G., et al. (2004). The thalamus as the generator and modulator of EEG alpha rhythm: a combined PET/EEG study with lorazepam challenge in humans. *Neuroimage* 22, 637–644. doi: 10.1016/j.neuroimage.2004.01.047
- Seth, A. K., Chorley, P., and Barnett, L. C. (2013). Granger causality analysis of fMRI BOLD signals is invariant to hemodynamic convolution but not downsampling. *Neuroimage* 65, 540–555. doi: 10.1016/j.neuroimage.2012.09.049
- Sheline, Y. I., Morris, J. C., Snyder, A. Z., Price, J. L., Yan, Z., D'Angelo, G., et al. (2010). APOE4 allele disrupts resting state fMRI connectivity in the absence of amyloid plaques or decreased CSF Aβ42. *J. Neurosci.* 30, 17035–17040. doi: 10.1523/JNEUROSCI.3987-10.2010
- Sheline, Y. I., Raichle, M. E., Snyder, A. Z., Morris, J. C., Head, D., Wang, S., et al. (2011). Amyloid plaques disrupt resting state default mode network connectivity in cognitively normal elderly. *Biol. Psychiatry* 67, 584–587. doi: 10.1016/j.biopsych.2009.08.024
- Sherman, S. M. (2007). The thalamus is more than just a relay. *Curr. Opin. Neurobiol.* 17, 417–422. doi: 10.1016/j.conb.2007.07.003
- Sherman, S. M., and Guillery, R. W. (2006). *Exploring the Thalamus*. San Diego, CA: Academic Press.
- Shiino, A., Watanabe, T., Maeda, K., Kotani, E., Akiguchi, I., and Matsuda, M. (2006). Four subgroups of Alzheimer's disease based on patterns of atrophy using VBM and a unique pattern for early onset disease. *Neuroimage* 33, 17–26. doi: 10.1016/j.neuroimage.2006.06.010
- Smith, S. M. (2002). Fast robust automated brain extraction. *Hum. Brain Mapp.* 17, 143–155. doi: 10.1002/hbm.10062
- Smith, S. M., Miller, K. L., Salimi-Khorshidi, G., Webster, M., Beckmann, C. F., Nichols, T. E., et al. (2011). Network modelling methods for FMRI. *Neuroimage* 54, 875–891. doi: 10.1016/j.neuroimage.2010.08.063
- Sorg, C., Riedel, V., Mühlau, M., Calhoun, V. D., Eichele, T., Läer, L., et al. (2007). Selective changes of resting-state networks in individuals at risk for Alzheimer's disease. *Proc. Natl. Acad. Sci. U S A* 104, 18760–18765. doi: 10.1073/pnas.0708803104
- Stephan, B. C. M., Hunter, S., Harris, D., Llewellyn, D. J., Siervo, M., Matthews, F. E., et al. (2012). The neuropathological profile of mild cognitive impairment (MCI): a systematic review. *Mol. Psychiatry* 17, 1056–1076. doi: 10.1038/mp.2011.147
- Teipel, S., and Grothe, M. J. (2016). Does posterior cingulate hypometabolism result from disconnection or local pathology across preclinical and clinical stages of Alzheimer's disease? *Eur. J. Nucl. Med. Mol. Imaging* 43, 526–536. doi: 10.1007/s00259-015-3222-3
- Thiebaut de Schotten, M., Dell'Acqua, F., Forkel, S. J., Simmons, A., Vergani, F., Murphy, D. G. M., et al. (2011). A lateralized brain network for visuospatial attention. *Nat. Neurosci.* 14, 1245–1246. doi: 10.1038/nn.2905
- Tournier, J. D., Calamante, F., and Connelly, A. (2007). Robust determination of the fibre orientation distribution in diffusion MRI: non-negativity constrained super-resolved spherical deconvolution. *Neuroimage* 35, 1459–1472. doi: 10.1016/j.neuroimage.2007.02.016
- Tournier, J.-D., Calamante, F., and Connelly, A. (2012). MRtrix: diffusion tractography in crossing fiber regions. *Int. J. Imaging Syst. Technol.* 22, 53–66. doi: 10.1002/ima.22005
- Tournier, J.-D., Calamante, F., Gadian, D. G., and Connelly, A. (2004). Direct estimation of the fiber orientation density function from diffusion-weighted MRI data using spherical deconvolution. *Neuroimage* 23, 1176–1185. doi: 10.1016/j.neuroimage.2004.07.037
- Tournier, J.-D., Yeh, C.-H., Calamante, F., Cho, K.-H., Connelly, A., and Lin, C.-P. (2008). Resolving crossing fibres using constrained spherical deconvolution: validation using diffusion-weighted imaging phantom data. *Neuroimage* 42, 617–625. doi: 10.1016/j.neuroimage.2008.05.002
- Valla, J., Berndt, J. D., and Gonzalez-Lima, F. (2001). Energy hypometabolism in posterior cingulate cortex of Alzheimer's patients: superficial laminar cytochrome oxidase associated with disease duration. *J. Neurosci.* 21, 4923–4930.
- Van Essen, D. C., Smith, S. M., Barch, D. M., Behrens, T. E. J., Yacoub, E., Ugurbil, K., et al. (2013). The WU-minn human connectome project: an overview. *Neuroimage* 80, 62–79. doi: 10.1016/j.neuroimage.2013.05.041
- Vann, S. D., Aggleton, J. P., and Maguire, E. A. (2009). What does the retrosplenial cortex do? *Nat. Rev. Neurosci.* 10, 792–802. doi: 10.1038/nrn2733
- Villain, N., Desgranges, B., Viader, F., de la Sayette, V., Mézenge, F., Landeau, B., et al. (2008). Relationships between hippocampal atrophy, white matter disruption and gray matter hypometabolism in Alzheimer's disease. *J. Neurosci.* 28, 6174–6181. doi: 10.1523/JNEUROSCI.1392-08.2008
- Wang, D., Kong, Y., Chu, W. C. W., Tam, C. W. C., Lam, L. C. W., Wang, Y., et al. (2014). Generation of the probabilistic template of default mode network derived from resting-state fMRI. *IEEE Trans. Biomed. Eng.* 61, 2550–2555. doi: 10.1109/TBME.2014.2323078
- Wang, L., Roe, C. M., Snyder, A. Z., Brier, M. R., Thomas, J. B., Xiong, C., et al. (2012). Alzheimer disease family history impacts resting state functional connectivity. *Ann. Neurol.* 72, 571–577. doi: 10.1002/ana.23643
- Welsh, K., Butters, N., Hughes, J., Mohs, R., and Heyman, A. (1991). Detection of abnormal memory decline in mild cases of Alzheimer's disease using CERAD neuropsychological measures. *Arch. Neurol.* 48, 278–281. doi: 10.1001/archneur.1991.00530150046016
- Welsh, K. A., Butters, N., Hughes, J. P., Mohs, R. C., and Heyman, A. (1992). Detection and staging of dementia in Alzheimer's disease. Use of the neuropsychological measures developed for the consortium to establish a registry for Alzheimer's disease. *Arch. Neurol.* 49, 448–452. doi: 10.1001/archneur.1992.00530290030008
- Wen, X., Rangarajan, G., and Ding, M. (2013). Is granger causality a viable technique for analyzing fMRI data? *PLoS One* 8:e67428. doi: 10.1371/journal.pone.0067428
- Wermke, M., Sorg, C., Wohlschläger, A. M., and Drzezga, A. (2008). A new integrative model of cerebral activation, deactivation and default mode

- function in Alzheimer's disease. *Eur. J. Nucl. Med. Mol. Imaging* 35, S12–S24. doi: 10.1007/s00259-007-0698-5
- Wheeler-Kingshott, C. A. M., and Cercignani, M. (2009). About “axial” and “radial” diffusivities. *Magn. Reson. Med.* 61, 1255–1260. doi: 10.1002/mrm.21965
- Wiener, N. (1956). “The theory of prediction,” in *Modern Mathematics for Engineers*, (Vol. 58) ed. E. F. Beckenbach (New York, NY: McGraw-Hill), 323–329.
- Xie, S., Zuo, N., Shang, L., Song, M., Fan, L., and Jiang, T. (2015). How does B-value affect HARDI reconstruction using clinical diffusion MRI data? *PLoS One* 10:e0120773. doi: 10.1371/journal.pone.0120773
- Xuereb, J. H., Perry, R. H., Candy, J. M., Perry, E. K., Marshall, E., and Bonham, J. R. (1991). Nerve cell loss in the thalamus in Alzheimer's disease and Parkinson's disease. *Brain* 114, 1363–1379. doi: 10.1093/brain/114.3.1363
- Yakushev, I., Schreckenberger, M., Müller, M. J., Schermuly, I., Cumming, P., Stoeter, P., et al. (2011). Functional implications of hippocampal degeneration in early Alzheimer's disease: a combined DTI and PET study. *Eur. J. Nucl. Med. Mol. Imaging* 38, 2219–2227. doi: 10.1007/s00259-011-1882-1
- Yeatman, J. D., Weiner, K. S., Pestilli, F., Rokem, A., Mezer, A., and Wandell, B. A. (2014). The vertical occipital fasciculus: a century of controversy resolved by *in vivo* measurements. *Proc. Natl. Acad. Sci. U S A* 111, E5214–E5223. doi: 10.1073/pnas.1418503111
- Yesavage, J. A., Brink, T. L., Rose, T. L., Lum, O., Huang, V., Adey, M., et al. (1983). Development and validation of a geriatric depression screening scale: a preliminary report. *J. Psychiatr. Res.* 17, 37–49. doi: 10.1016/0022-3956(82)90033-4
- Zarei, M., Beckmann, C. F., Binnewijzend, M. A. A., Schoonheim, M. M., Oghabian, M. A., Sanz-Arigita, E. J., et al. (2013). Functional segmentation of the hippocampus in the healthy human brain and in Alzheimer's disease. *Neuroimage* 66, 28–35. doi: 10.1016/j.neuroimage.2012.10.071
- Zarei, M., Patenaude, B., Damoiseaux, J., Morgese, C., Smith, S., Matthews, P. M., et al. (2010). Combining shape and connectivity analysis: an MRI study of thalamic degeneration in Alzheimer's disease. *Neuroimage* 49, 1–8. doi: 10.1016/j.neuroimage.2009.09.001
- Zhang, Y., Brady, M., and Smith, S. (2001). Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. *IEEE Trans. Med. Imaging* 20, 45–57. doi: 10.1109/42.906424
- Zhang, Y., Schuff, N., Camacho, M., Chao, L. L., Fletcher, T. P., Yaffe, K., et al. (2013). MRI markers for mild cognitive impairment: comparisons between white matter integrity and gray matter volume measurements. *PLoS One* 8:e66367. doi: 10.1371/journal.pone.0066367
- Zhang, D., Snyder, A. Z., Fox, M. D., Sansbury, M. W., Shimony, J. S., and Raichle, M. E. (2008). Intrinsic functional relations between human cerebral cortex and thalamus. *J. Neurophysiol.* 100, 1740–1748. doi: 10.1152/jn.90463.2008
- Zhang, D., Snyder, A. Z., Shimony, J. S., Fox, M. D., and Raichle, M. E. (2010). Noninvasive functional and structural connectivity mapping of the human thalamocortical system. *Cereb. Cortex* 20, 1187–1194. doi: 10.1093/cercor/bhp182
- Zhu, D. C., Majumdar, S., Korolev, I. O., Berger, K. L., and Bozoki, A. C. (2013). Alzheimer's disease and amnesic mild cognitive impairment weaken connections within the default-mode network: a multi-modal imaging study. *J. Alzheimers Dis.* 34, 969–984. doi: 10.3233/JAD-121879

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.

Copyright © 2017 Alderson, Kehoe, Maguire, Farrell, Lawlor, Kenny, Lyons, Bokde and Coyle. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Preprocessing of ^{18}F -DMFP-PET Data Based on Hidden Markov Random Fields and the Gaussian Distribution

Fermin Segovia^{1*}, Juan M. Górriz^{1,2}, Javier Ramírez¹, Francisco J. Martínez-Murcia¹ and Diego Salas-Gonzalez¹

¹ Department of Signal Theory, Networking and Communications, University of Granada, Granada, Spain, ² Department of Psychiatry, University of Cambridge, Cambridge, United Kingdom

OPEN ACCESS

Edited by:

Pedro Rosa-Neto,
McGill University, Canada

Reviewed by:

Manuel Grana,
University of the Basque Country
(UPV/EHU), Spain
Ana Maria Tome,
University of Aveiro, Portugal

*Correspondence:

Fermin Segovia
fsegovia@ugr.es

Received: 18 March 2017

Accepted: 20 September 2017

Published: 09 October 2017

Citation:

Segovia F, Górriz JM, Ramírez J, Martínez-Murcia FJ and Salas-Gonzalez D (2017) Preprocessing of ^{18}F -DMFP-PET Data Based on Hidden Markov Random Fields and the Gaussian Distribution. *Front. Aging Neurosci.* 9:326. doi: 10.3389/fnagi.2017.00326

^{18}F -DMFP-PET is an emerging neuroimaging modality used to diagnose Parkinson's disease (PD) that allows us to examine postsynaptic dopamine $D_{2/3}$ receptors. Like other neuroimaging modalities used for PD diagnosis, most of the total intensity of ^{18}F -DMFP-PET images is concentrated in the striatum. However, other regions can also be useful for diagnostic purposes. An appropriate delimitation of the regions of interest contained in ^{18}F -DMFP-PET data is crucial to improve the automatic diagnosis of PD. In this manuscript we propose a novel methodology to preprocess ^{18}F -DMFP-PET data that improves the accuracy of computer aided diagnosis systems for PD. First, the data were segmented using an algorithm based on Hidden Markov Random Field. As a result, each neuroimage was divided into 4 maps according to the intensity and the neighborhood of the voxels. The maps were then individually normalized so that the shape of their histograms could be modeled by a Gaussian distribution with equal parameters for all the neuroimages. This approach was evaluated using a dataset with neuroimaging data from 87 parkinsonian patients. After these preprocessing steps, a Support Vector Machine classifier was used to separate idiopathic and non-idiopathic PD. Data preprocessed by the proposed method provided higher accuracy results than the ones preprocessed with previous approaches.

Keywords: PET image segmentation, ^{18}F -DMFP-PET data, intensity normalization, Hidden Markov Models, Gaussian distribution, Parkinson's disease

1. INTRODUCTION

Neuroimaging data have become an essential tool to diagnose the most frequent neurodegenerative disorders: Alzheimer's and Parkinson's disease. Initially, the neuroimages were visually inspected by experienced clinicians in order to corroborate a previous tentative diagnosis based on neuropsychological and behavioral tests. To this end, they looked for areas of low activation located in specific brain regions that are known to be affected by the supposed disorder. During the last decade, the neuroimaging community has progressively increased the use of computer toolboxes to analyze neuroimaging data (Friston et al., 2006; Schrouff et al., 2013). These tools are able to carry out statistical analyses that perform a more exhaustive examination of the huge amounts of information contained in the data and remove the subjectivity inherent to the visual inspection

of the neuroimages (Górriz et al., 2017). However, these statistical analyses require additional preprocessing steps that make neuroimages from different subjects comparable. Two procedures are usually performed: spatial registration and intensity normalization (Saxena et al., 1998; Dukart et al., 2013). The former ensures that a given voxel from different neuroimages corresponds to the same anatomical position while the latter removes the differences due to the scanner used or the amount of radiopharmaceutical injected (Salas-Gonzalez et al., 2013). Even when all the data are acquired using a single scanner, an intensity normalization of the data is desirable. Several studies suggested that the absolute values of cerebral blood flow and other metabolic measurements have a coefficient of variance about 15% in healthy elderly subjects and as high as 30% in patients suffering neurodegenerative disorders (Leenders et al., 1990; Huang et al., 2007; Borghammer et al., 2009). In addition to spatial registration and intensity normalization, a segmentation step can be carried out. This procedure consists on partitioning the data into two or more maps each one containing information of different classes. For example, brain Magnetic Resonance Imaging (MRI) data are usually segmented into gray matter, white matter and cerebrospinal fluid. Segmentation is common in studies that use structural data but it has been also used for functional data. In Moussallem et al. (2012) the authors used a threshold (adjusted by an *ad hoc* function) to segment ^{18}F -FDG-PET data in order to delimit tumors. A more sophisticated approach for the same purpose was demonstrated in Li et al. (2017).

^{18}F -DMFP-PET is a neuroimage modality that is increasingly being used as an effective tool to distinguish between idiopathic and non-idiopathic parkinsonian patients and therefore to assist the diagnosis of Parkinson's disease (PD) (la Fougère et al., 2010). In contrast to DaTSCAN (widely used for PD diagnosis; Towey et al., 2011; Illán et al., 2012; Segovia et al., 2012; Martínez-Murcia et al., 2014), ^{18}F -DMFP-PET is able to image the postsynaptic striatal dopaminergic deficit that characterizes non-idiopathic parkinsonian variants such as multiple system atrophy (MSA) or progressive supranuclear palsy (PSP). Because of this, most of the studies with ^{18}F -DMFP-PET are focused on analyzing the striatal region, even though this neuroimaging modality contains moderate signal intensities in regions other than the striatum that can be useful in PD diagnosis (Segovia et al., 2015, 2017).

In this work, we propose a methodology to preprocess ^{18}F -DMFP-PET data that improves the results of subsequent analyses. The proposed method consists of two steps: data segmentation using Hidden Markov Random Fields (HMRF) and intensity normalization using the Gaussian distribution. The segmentation step divides each neuroimage into 4 maps: (i) high-signal voxels (located in the striatum), (ii) medium-signal voxels (located in most of the regions other than the striatum), (iii) low-signal voxels (most of them correspond to the cerebrospinal fluid), and (iv) voxels with intensities around zero (located outside the brain). The second step normalizes the intensities of each map using a Gaussian model. This approach was evaluated and compared with previous approaches using 87 neuroimages and a system based on Support Vector Machine (SVM) classification (Vapnik, 2000). The obtained results suggest that our procedure

improves the automatic separation of idiopathic and non-idiopathic parkinsonian patients. In addition, it allows us to independently analyze the striatum and the remaining regions of the brain.

2. MATERIALS AND METHODS

2.1. Ethics Statement

Each patient (or a close relative) gave written informed consent to participate in the study and the protocol was accepted by the Ethics Committee of the University of Munich. All the data were anonymized by the clinicians who acquired them before being considered in this work.

2.2. ^{18}F -DMFP-PET Neuroimaging Database

Eighty-seven ^{18}F -DMFP-PET neuroimages were used to evaluate the preprocessing approach proposed in this work. These data were collected during a longitudinal study carried out by the University of Munich (la Fougère et al., 2010). The neuroimages were acquired 55 min after the ^{18}F -DMFP injection (which was synthesized using an automatic synthesis module as described in la Fougère et al., 2010) by means of a Siemens/CTI camera. Neuroleptics, metoclopramide and other medications and dopamine agonists that could potentially interfere were withdrawn before the data acquisition according to their biologic half-life. The emission recording consisted of 3 frames of 10 min each, acquired in 3-dimensional mode. The resulting images were reconstructed as 128×128 matrices of 2×2 mm voxels by filtered backprojection using a Hann filter.

All the patients included in the study were referred to ^{18}F -DMFP-PET examination from local movement disorder clinics. They showed parkinsonian movement disorders and nigrostriatal degeneration that were confirmed by a ^{123}I -FP-CIT SPECT scan according to widely accepted criteria (Koch et al., 2005). They were monitored during 2 years after the ^{18}F -DMFP-PET acquisition and at this time the neuroimaging data were labeled according to the last diagnosis. Specifically, the last diagnosis was based on the response to an apomorphine challenge test or the response to dopamine replacement therapy and follow-up clinical examinations, paying special attention to orthostatic hypotension, cerebellar signs, eye movement disorders, spasticity or other atypical symptoms. **Table 1** shows the resulting groups and some demographic details.

TABLE 1 | Group distribution of the neuroimaging data considered in this work (μ and σ stand for the mean and the standard deviation, respectively).

| | # | Sex | | Age | | Range |
|-----|----|-----|----|-------|----------|-------|
| | | M | F | μ | σ | |
| PD | 39 | 22 | 17 | 61.38 | 11.14 | 35–81 |
| MSA | 24 | 20 | 4 | 68.42 | 10.73 | 43–85 |
| PSP | 24 | 12 | 12 | 69.29 | 7.33 | 55–84 |

Before applying the proposed method and the subsequent classification, the data were spatially registered using the template matching algorithm implemented in Statistical Parametric Mapping (SPM) (Friston et al., 2006). This procedure makes each neuroimage matches a given template, pursuing the same position (in the neuroimage space) in different neuroimages corresponds to the same anatomical position. The template was computed as follows: first all the neuroimages were registered to a randomly chosen one. The registered images and their hemisphere midplane reflections were then averaged (this step ensured a symmetric template). Finally the resulting image was smoothed and used to register the whole dataset (Ashburner et al., 1997).

2.3. Markov Models

A Markov model (a.k.a. Markov chain) is a discrete stochastic process in which the next state only depends on the current state. If unobserved (hidden) states are assumed, the model is known as hidden Markov model (HMM). This work is focused on Markov random fields (MRF) that can be considered a generalization of Markov models for multiple-dimensions problems.

2.3.1. Markov Random Fields

Markov random field theory is a branch of probability theory for analyzing the spatial or contextual dependencies of physical phenomena. A MRF is a family of random variables that satisfies the Markovianity property and can be described by an undirected graphical model.

Let $S = \{1, 2, \dots, N\}$ be the set of indexes in space, and $\mathcal{N} = \{\mathcal{N}_i, i \in S\}$ a neighborhood system, with \mathcal{N}_i being the set of sites neighboring i and satisfying that $i \notin \mathcal{N}_i$ and $i \in \mathcal{N}_j \iff j \in \mathcal{N}_i$. A random field is said to be a MRF on S with respect to a neighborhood system \mathcal{N} if and only if Li (2001):

$$\begin{aligned} P(\mathbf{x}) &> 0, \forall \mathbf{x} \in \chi \\ P(x_i | x_{S-\{i\}}) &= P(x_i | \mathcal{N}_i) \end{aligned} \quad (1)$$

where $\mathbf{x} = (x_1, x_2, \dots, x_N)$ is a configuration in S and χ is the set of all possible configurations in S . A MRF can be characterized by a Gibbs distribution, allowing us to redefine the probability $P(\mathbf{x})$ as (Hammersley-Clifford theorem):

$$P(\mathbf{x}) = \frac{1}{Z} \exp(-T^{-1}U(\mathbf{x})) \quad (2)$$

where:

$$Z = \sum_{\mathbf{x} \in \chi} \exp(-T^{-1}U(\mathbf{x})) \quad (3)$$

is a normalizing constant, T is a constant called temperature and usually fixed to 1 and $U(\mathbf{x})$ is the energy function, defined as a sum of clique potentials $V_c(\mathbf{x})$ over all possible cliques, C :

$$U(\mathbf{x}) = \sum_{c \in C} V_c(\mathbf{x}) \quad (4)$$

In this context, a clique c for the graph constituted by S and \mathcal{N} (S contains the nodes and \mathcal{N} the links) is defined as a subset of S whose elements are neighbors to one another (Li, 2001).

2.3.2. Hidden Markow Random Fields

Hidden Markow random fields are a generalization of HMMs that assume MRFs (more than one dimension) instead of Markov models (one dimension) and therefore, they can be directly applied to two and three-dimensional problems, such as neuroimage segmentation.

A HMRF is characterized by an unobservable (hidden) MRF $X = \{X_i, i \in S\}$ assuming values in a finite state space L , an observable random field $Y = \{Y_i, i \in S\}$ assuming values in a finite state space D , and a conditional independence restriction (Zhang et al., 2001). For any particular configuration $\mathbf{x} \in \chi$, every Y_i follows a known conditional probability distribution $p(y_i | x_i)$ of the same functional form $f(y_i; \theta_{x_i})$. Given that (conditional independence):

$$P(\mathbf{y} | \mathbf{x}) = \prod_{i \in S} P(y_i | x_i) \quad (5)$$

the joint probability of (X, Y) can be written as:

$$P(\mathbf{y}, \mathbf{x}) = P(\mathbf{y} | \mathbf{x})P(\mathbf{x}) = P(\mathbf{x}) \prod_{i \in S} P(y_i | x_i) \quad (6)$$

Since $P(y_i, x_i | x_{\mathcal{N}_i}) = P(y_i | x_i)P(x_i | x_{\mathcal{N}_i})$ (because of the local characteristics of MRFs), the marginal probability distribution of Y_i can be computed in function of the parameter set θ and $X_{\mathcal{N}_i}$:

$$p(y_i | x_{\mathcal{N}_i}, \theta) = \sum_{l \in L} p(y_i, l | x_{\mathcal{N}_i}, \theta) = \sum_{l \in L} f(y_i; \theta_l) p(l | x_{\mathcal{N}_i}) \quad (7)$$

2.4. Automatic Segmentation Based On HMRF

The segmentation based on HMRF assigns a label $l_i \in L = \{1, 2, 3, 4\}$, $i = \{1, \dots, N\}$ to each voxel in a ^{18}F -DMFP-PET neuroimage according to both intensity and neighborhood. Let $\mathbf{y} = \{y_1, y_2, \dots, y_N\}$ be the intensity levels of the N voxels that form a ^{18}F -DMFP-PET neuroimage. In this procedure we looked for a labeling $\mathbf{x} = (x_1, x_2, \dots, x_N)$, where $x_i \in L$ is the label assigned to the voxel y_i . Formally we estimated (MAP criterion):

$$\hat{\mathbf{x}} = \arg \max_{\mathbf{x} \in \chi} \{P(\mathbf{y} | \mathbf{x})P(\mathbf{x})\} \quad (8)$$

where $\hat{\mathbf{x}}$ is an estimation of \mathbf{x} and considered a particular realization of the MRF X . Using the equivalence between MRFs and Gibb distributions the Equation (8) can be written as Zhang et al. (2001):

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x} \in \chi} \{U(\mathbf{y} | \mathbf{x}) + U(\mathbf{x})\} \quad (9)$$

where $U(\mathbf{y}|\mathbf{x})$ is the likelihood energy. Estimating $\hat{\mathbf{x}}$ involves estimating the parameter set $\theta = \{\theta_l, l \in L\}$, where $\theta_l = (\mu_l, \sigma_l)$, since we assumed a Gaussian function for each of the maps resulting from the segmentation of \mathbf{y} . A k -means algorithm was used to initialize the labeling $\hat{\mathbf{x}}$. Then, an Expectation-Maximization (EM) algorithm was carried out to alternatively estimate the parameter set, θ , and the label set, $\hat{\mathbf{x}}$.

Altogether, this segmentation procedure divides a neuroimage into 4 maps: (i) voxels with intensity close to zero (mainly voxels outside the brain), (ii) low-signal voxels, with very limited diagnostic value, (iii) medium-signal voxels, and (iv) high-signal voxels, with high diagnostic value (located in the striatal region). In order to reduce the computational burden, the segmentation procedure was only applied to an *ad-hoc* neuroimage computed as the average of all the ^{18}F -DMFP-PET images in our dataset (the result is shown in **Figure 1**). Then, the resulting maps were used as binary masks to segment the neuroimages in our dataset.

In this initial work, only striatal voxels were considered to separate idiopathic and non-idiopathic patients. Thus, only the maps containing high-signal voxels (one map per neuroimage) were used in the subsequent analyses.

2.5. Intensity Normalization Based on the Gaussian Distribution

The intensity of high-signal voxels largely differs from one patient to another, even among patients suffering the same

parkinsonian disorder. This can be noted on **Figure 2**, which shows the histogram of the map containing these voxels for the first 20 patients in our dataset (all of them were diagnosed with idiopathic parkinsonism).

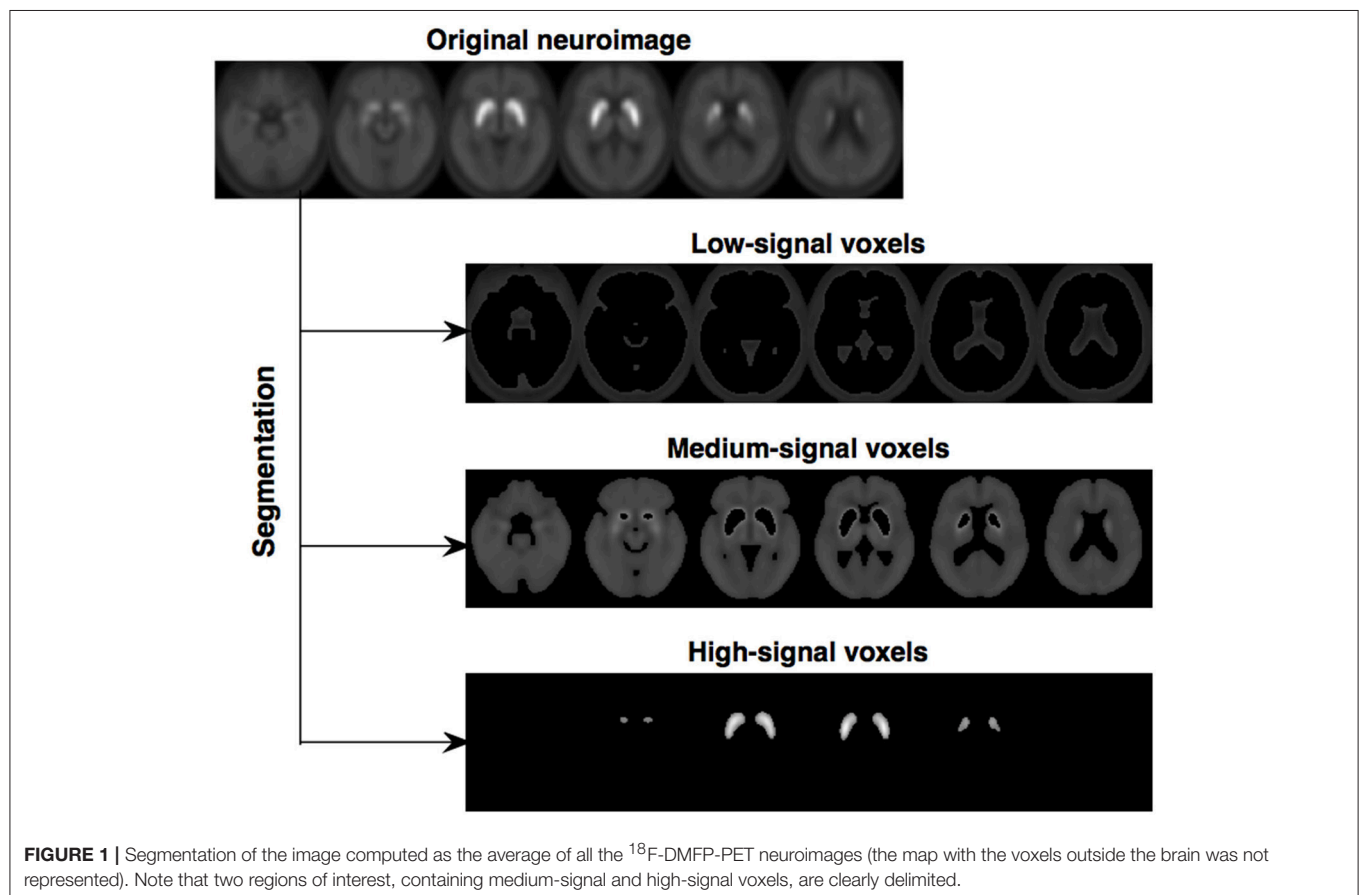
In order to reduce these differences without losing the discriminant information contained in the data, an additional normalization step was performed. This procedure modeled the histogram of a given map of each patient by a Gaussian distribution. Then, these data were modified so that the Gaussians corresponding to all the patients have approximately same mean and standard deviation. First, parameters G_μ and G_σ were computed:

$$G_\mu = \frac{1}{n} \sum_{i=1}^n \mu_{\mathbf{p}_i} \quad (10)$$

$$G_\sigma = \frac{1}{n} \sum_{i=1}^n \sigma_{\mathbf{p}_i} \quad (11)$$

where $\mu_{\mathbf{p}_i}$ and $\sigma_{\mathbf{p}_i}$ respectively stand for the mean and standard deviation of the Gaussian associated to data from patient \mathbf{p}_i , and n is number of patients/neuroimages in our dataset. The data from each patient were then modified as follows:

$$\mathbf{p}_i^{(NORM)} = G_\sigma \frac{\mathbf{p}_i - \mu_{\mathbf{p}_i}}{\sigma_{\mathbf{p}_i}} + G_\mu \quad (12)$$



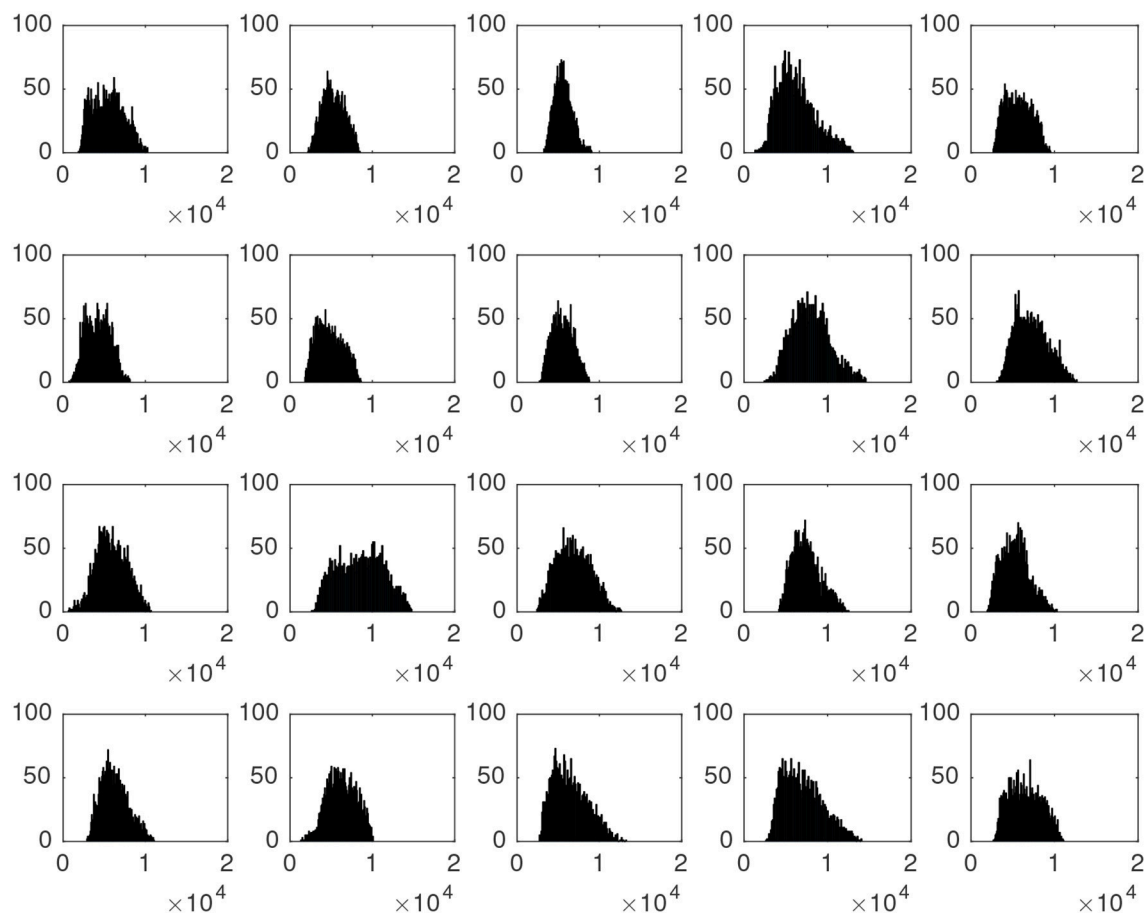


FIGURE 2 | Histograms associated to the striatal voxels of the first 20 neuroimages in our dataset. All of them correspond to idiopathic parkinsonian patients.

Figure 3 illustrates the transformation carried out by this procedure. It shows the shape of the histograms of our data before and after the normalization. Note that this procedure was independently applied to the data of each patient and in our case, it was only used to normalize the maps with high-signal voxels, however it can be also applied other maps obtained from the segmentation.

3. RESULTS AND DISCUSSION

In order to evaluate the advantages of preprocessing ^{18}F -DMFP-PET data with our methodology, a statistical classification analysis was carried out. To this purpose, a SVM classifier (Vapnik, 2000) was used after the preprocessing steps to separate the idiopathic and non-idiopathic patients in our database (i.e., PD vs. MSA and PSP). As it is common in PD diagnosis (Winogrodzka et al., 2001; Constantinescu et al., 2011; Niccolini et al., 2014; Prashanth et al., 2014) we used only the voxels at the striatum, as selected by the maps with high-signal voxels (the other maps resulting from the segmentation were discarded). The normalized intensity values of the selected voxels were directly used as feature.

The classification performance was estimated by means of a k -fold cross-validation scheme ($k = 5$). In order to avoid biased results, all the parameters required by the method were fit inside the cross-validation loop, using only the training data. A nested loop was also used to adjust the parameter C of the SVM classifier (Varma and Simon, 2006). **Table 2** shows the achieved accuracy, sensitivity and specificity (idiopathic patients were considered as positive) and compares these results with the ones obtained by other approaches: (i) selecting voxels at the striatum by means of an atlas and, (ii) using all the voxels of the brain. In these cases, the intensity of the voxels was normalized using the normalization to the maximum (Saxena et al., 1998).

The results shown in **Table 2** suggest that our preprocessing method allows improving the automatic separation of parkinsonian patients. The relatively low rates achieved by the SVM classifier are due to the dataset used in this work. Most of the neuroimages correspond to patients in a very initial stage. In fact they were acquired 2 years before obtaining the final diagnosis used to label the data. In addition, the whole brain approach also suffers from the small sample size problem (Duin, 2000). In this classification the number of features is larger and many of these features correspond to regions of low signal in ^{18}F -DMFP-PET data, which are not useful to separate

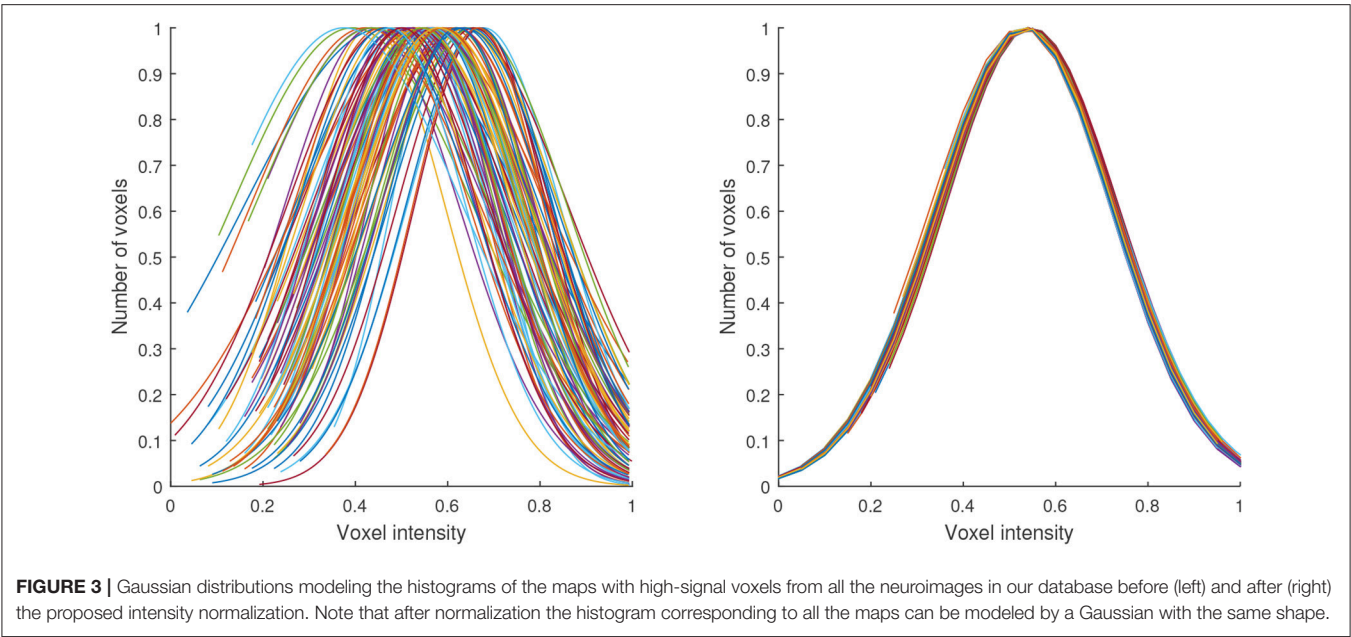


TABLE 2 | Accuracy, sensitivity and specificity obtained by a SVM classifier when separating idiopathic and non-idiopathic Parkinsonism.

| Features set | Accuracy (%) | Sensitivity (%) | Specificity (%) |
|----------------------------|--------------|-----------------|-----------------|
| Striatum (proposed method) | 75.86 | 74.36 | 77.08 |
| Striatum (atlas) | 72.41 | 66.67 | 77.08 |
| All the voxels | 65.52 | 56.41 | 72.92 |

the groups. In terms of sensitivity and specificity, the obtained results show that the proposed method largely improve (about 8%) the ability of the classifier to correctly detect the positive subjects (idiopathic Parkinsonism) however the improvement in the true negative rate is limited, specially when compared with the atlas-based approach. This fact can be explained by the heterogeneity of the negative group (composed by subjects diagnosed with MSA and PSP), which makes more difficult to characterize the data.

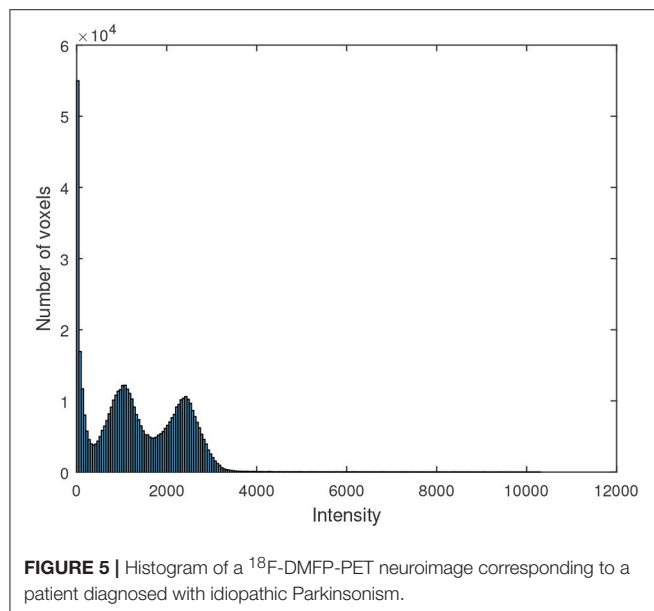
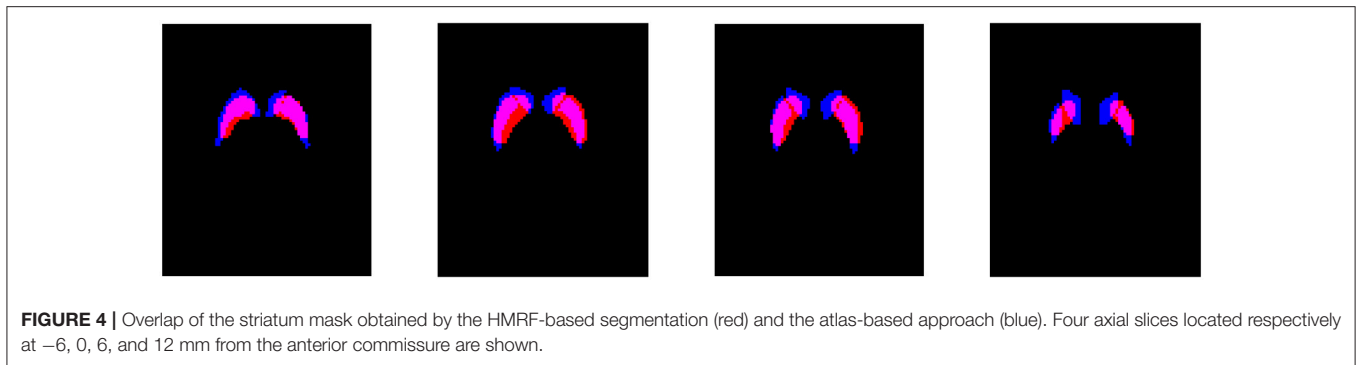
As mentioned above, the analysis of neuroimaging data for diagnostic purposes in PD-related studies is commonly focused on the striatum. In fact, post-mortem studies revealed that most of the neuropathological hallmarks of PD are gathered in this area (Rinne et al., 1991; Hartmann, 2004; Nagatsu and Sawada, 2007). Nonetheless the region to be analyzed highly depends on the neuroimage modality or, more specifically, on the binding properties of the radiotracer used. For example, studies using ¹²³I-FP-CIT (Winogrodzka et al., 2001; Spiegel et al., 2007) frequently constraint their analyses to the striatum, since this radiotracer binds to dopamine transporters, whereas studies based on ¹⁸F-FDG usually analyze the whole brain (Hellwig et al., 2012; Garraux et al., 2013) since this drug measures the brain metabolism. ¹⁸F-DMFP is commonly used to study the striatal dopamine (Schreckenberger et al., 2004) and indeed, the

vast majority of high-intensity voxels in ¹⁸F-DMFP-PET images are gathered in the striatum. However, these data show a not insignificant part of the total intensity in regions other than the striatum (Segovia et al., 2017). The segmentation methodology proposed in this work allows scientists to independently analyze high-signal and medium-signal voxels (respectively located in the striatum and in the remaining regions in ¹⁸F-DMFP-PET data) while low-signal voxels (with low signal-noise ratio) are discarded.

Compared with an atlas-based approach, our segmentation method not only provides a higher accuracy in the subsequent classification procedure but also allows the separation of regions of interest in the image space. Thus, it is not necessary to transform the data to the atlas space, avoiding the distortions introduced by these procedures (Ashburner and Friston, 2007).

A comparison between the striatum region obtained by the HMRF-based segmentation method and the atlas-based approach is shown in **Figure 4**. A quantitative analysis of this comparison reveals that: (i) the striatum region is about 30% larger when obtained by means of the atlas-based approach; (ii) most of the voxels selected by the proposed method (about 72%) were also selected by the other approach. Thus, the improvements in the classification procedure are probably because the HMRF-based segmentation provide a more accurate delimitation of the discriminant voxels. Most of these voxels are located in the striatum but not all the voxels in the striatum should be considered to separate idiopathic and non-idiopathic Parkinsonism. According to the results shown in **Table 2**, discarding these moderately discriminant voxels of the striatum provides larger sensitivity rates but have a reduced impact in the specificity.

The motivation to use Gaussian distributions to model the histogram of the maps resulting from the segmentation is



explained by the **Figure 5**, which shows the histogram of a ^{18}F -DMFP-PET neuroimage. The two Gaussians corresponding to maps with the low-signal and medium-signal-voxels can be clearly identified. The Gaussian for high-signal voxels has much less height than the remaining ones and can not be appreciated in **Figure 5** but it can be identified in the histograms of **Figure 2**. Finally, the voxels with intensity very close to zero could be modeled by a fourth Gaussian. Indeed, the segmentation of a ^{18}F -DMFP-PET neuroimage using this algorithm is similar to model the histogram of that neuroimage by a sum (or mixture) of 4 Gaussians (Segovia et al., 2010; Górriz et al., 2011). Nevertheless, the HMRF approach takes into account both, the voxel intensity and the voxel neighborhood to associate each voxel to a specific map/Gaussian.

In this work, the segmentation method was applied only to an average neuroimage and the result was used to parcel each individual neuroimage. This approach requires lower computational burden than the straightforward alternative consisting on applying the segmentation algorithm to each neuroimage. Additionally the resulting maps are of equal size for

all the neuroimages, what allows us to directly use the voxels as feature in the subsequent classification step.

4. CONCLUSION

In this manuscript we described a novel methodology to preprocess ^{18}F -DMFP-PET data in order to improve the diagnosis of Parkinsonism. The preprocessing method was carried out in two steps. First, using a HMRF-based approach, each neuroimage was divided into 4 maps according to the intensity and the neighborhood of the voxels. Then, the intensity of the voxels was normalized using the properties of the Gaussian distribution. To this end, the histogram of each map was modeled by a Gaussian distribution with the same parameters for all the neuroimages.

This methodology was evaluated using a dataset with neuroimaging data from 87 patients diagnosed with idiopathic or non-idiopathic Parkinsonism. Using the proposed methodology, we selected and normalized the high-signal voxels of each neuroimage. These data were used to train a SVM classifier in order to separate idiopathic and non-idiopathic subjects, obtaining an accuracy rate about 75%. These results outperform those reported by previous approaches, what suggests that our preprocessing method improves the computer tools currently used to assist the diagnosis of Parkinsonism.

AUTHOR CONTRIBUTIONS

Drafting the article and conception or design of the work: FS. Critical revision of the article, data analysis and interpretation: FS, JG, JR, FM, and DS.

ACKNOWLEDGMENTS

The authors are grateful to Johannes Levin, Axel Rominger, Madeleine Schuberth, and Matthias Brendel from the University of Munich for their help in data management. This work was supported by and the MINECO under the TEC2012-34306 and TEC2015-64718-R projects and the Ministry of Economy, Innovation, Science and Employment of the Junta de Andalucía under the Excellence Projects P09-TIC-4530 and

P11-TIC-7103 and a Talent Hub project approved by the Andalucía Talent Hub Program launched by the Andalusian Knowledge Agency, co-funded by the European Union's Seventh Framework Program, Marie Skłodowska-Curie actions (COFUND — Grant Agreement no. 291780) and the Ministry

of Economy, Innovation, Science and Employment of the Junta de Andalucía. The work was also supported by the Vicerectorate of Research and Knowledge Transfer of the University of Granada and the Salvador de Madariaga Mobility Grants 2017.

REFERENCES

- Ashburner, J., and Friston, K. (2007). "Segmentation," in *Statistical Parametric Mapping*, eds K. Friston, J. Ashburner, S. Kiebel, T. Nichols, and W. Penny (London: Academic Press), 81–91. doi: 10.1016/B978-012372560-8/50006-1
- Ashburner, J., Neelin, P., Collins, D. L., Evans, A., and Friston, K. (1997). Incorporating prior knowledge into image registration. *NeuroImage* 6, 344–352. doi: 10.1006/nimg.1997.0299
- Borghammer, P., Aanerud, J., and Gjedde, A. (2009). Data-driven intensity normalization of PET group comparison studies is superior to global mean normalization. *NeuroImage* 46, 981–988. doi: 10.1016/j.neuroimage.2009.03.021
- Constantinescu, C. C., Coleman, R. A., Pan, M.-L., and Mukherjee, J. (2011). Striatal and extrastriatal microPET imaging of D2/D3 dopamine receptors in rat brain with [¹⁸F]fallypride and [¹⁸F]desmethoxyfallypride. *Synapse (New York, N.Y.)* 65, 778–787. doi: 10.1002/syn.20904
- Duin, R. (2000). "Classifiers in almost empty spaces," in *15th International Conference on Pattern Recognition, 2000. Proceedings*, Vol. 2 (Barcelona), 1–7.
- Dukart, J., Perneczky, R., Förster, S., Barthel, H., Diehl-Schmid, J., Draganski, B., et al. (2013). Reference cluster normalization improves detection of frontotemporal lobar degeneration by means of FDG-PET. *PLoS ONE* 8:e55415. doi: 10.1371/journal.pone.0055415
- Friston, K. J., Ashburner, J. T., Kiebel, S. J., Nichols, T. E., and Penny, W. D. (2006). *Statistical Parametric Mapping: The Analysis of Functional Brain Images, 1st Edn*. Amsterdam; Boston: Academic Press.
- Garraux, G., Phillips, C., Schrouff, J., Kreisler, A., Lemaire, C., Degueldre, C., et al. (2013). Multiclass classification of FDG PET scans for the distinction between Parkinson's disease and atypical parkinsonian syndromes. *NeuroImage* 2, 883–893. doi: 10.1016/j.nicl.2013.06.004
- Górriz, J. M., Ramírez, J., Suckling, J., Illán, I. A., Ortiz, A., Martínez-Murcia, F. J., et al. (2017). Case-based statistical learning: a non-parametric implementation with a conditional-error rate SVM. *IEEE Access* 5, 11468–11478. doi: 10.1109/ACCESS.2017.2714579
- Górriz, J. M., Segovia, F., Ramírez, J., Lassl, A., and Salas-Gonzalez, D. (2011). GMM based SPECT image classification for the diagnosis of Alzheimer's disease. *Appl. Soft Comput.* 11, 2313–2325. doi: 10.1016/j.asoc.2010.08.012
- Hartmann, A. (2004). Postmortem studies in Parkinson's disease. *Dialogues Clin. Neurosci.* 6, 281–293.
- Hellwig, S., Amtage, F., Kreft, A., Buchert, R., Winz, O. H., Vach, W., et al. (2012). 18F-FDG-PET is superior to 123I-IBZM-SPECT for the differential diagnosis of parkinsonism. *Neurology* 79, 1314–1322. doi: 10.1212/WNL.0b013e31826c1b0a
- Huang, C., Tang, C., Feigin, A., Lesser, M., Ma, Y., Pourfar, M., et al. (2007). Changes in network activity with the progression of Parkinson's disease. *Brain* 130, 1834–1846. doi: 10.1093/brain/awm086
- Illán, I. A., Górriz, J. M., Ramírez, J., Segovia, F., Jiménez-Hoyuela, J. M., and Lozano, S. J. O. (2012). Automatic assistance to Parkinson's disease diagnosis in DaTSCAN SPECT imaging. *Med. Phys.* 39, 5971–5980. doi: 10.1118/1.4742055
- Koch, W., Radau, P. E., Hamann, C., and Tatsch, K. (2005). Clinical testing of an optimized software solution for an automated, observer-independent evaluation of dopamine transporter SPECT studies. *J. Nucl. Med.* 46, 1109–1118.
- la Fougère, C., Pöpperl, G., Levin, J., Wängler, B., Böning, G., Uebles, C., et al. (2010). The value of the dopamine D_{2/3} receptor ligand 18F-desmethoxyfallypride for the differentiation of idiopathic and nonidiopathic parkinsonian syndromes. *J. Nucl. Med.* 51, 581–587. doi: 10.2967/jnumed.109.071811
- Leenders, K., Perani, D., Lammertsma, A., Heather, J., Buckingham, P., Jones, T., et al. (1990). Cerebral blood flow, blood volume and oxygen utilization: normal values and effect of age. *Brain* 113, 27–47. doi: 10.1093/brain/113.1.27
- Li, L., Wang, J., Lu, W., and Tan, S. (2017). Simultaneous tumor segmentation, image restoration, and blur kernel estimation in PET using multiple regularizations. *Comput. Vis. Image Understand.* 155, 173–194. doi: 10.1016/j.cviu.2016.10.002
- Li, S. Z. (2001). *Markov Random Field Modeling in Image Analysis, 2 Edn*. Computer Science Workbench. Tokyo: Springer. doi: 10.1007/978-4-431-67044-5
- Martínez-Murcia, F. J., Górriz, J. M., Ramírez, J., Illán, I. A., and Ortiz, A. (2014). Automatic detection of Parkinsonism using significance measures and component analysis in DaTSCAN imaging. *Neurocomputing* 126, 58–70. doi: 10.1016/j.neucom.2013.01.054
- Moussallem, M., Valette, P.-J., Traverse-Glehen, A., Houzard, C., Jegou, C., and Giammarile, F. (2012). New strategy for automatic tumor segmentation by adaptive thresholding on PET/CT images. *J. Appl. Clin. Med. Phys.* 13:3875. doi: 10.1120/jacmp.v13i5.3875
- Nagatsu, T., and Sawada, M. (2007). Biochemistry of postmortem brains in Parkinson's disease: historical overview and future prospects. *J. Neural Transm. Suppl.* 72, 113–120. doi: 10.1007/978-3-211-73574-9_14
- Niccolini, F., Su, P., and Politis, M. (2014). Dopamine receptor mapping with PET imaging in Parkinson's disease. *J. Neurol.* 261, 2251–2263. doi: 10.1007/s00415-014-7302-2
- Prashanth, R., Dutta Roy, S., Mandal, P. K., and Ghosh, S. (2014). Automatic classification and prediction models for early Parkinson's disease diagnosis from SPECT imaging. *Exp. Syst. Appl.* 41, 3333–3342. doi: 10.1016/j.eswa.2013.11.031
- Rinne, J. O., Laihin, A., Lönnberg, P., Marjamäki, P., and Rinne, U. K. (1991). A post-mortem study on striatal dopamine receptors in Parkinson's disease. *Brain Res.* 556, 117–122. doi: 10.1016/0006-8993(91)90554-9
- Salas-Gonzalez, D., Górriz, J. M., Ramírez, J., Illán, I. A., and Lang, E. W. (2013). Linear intensity normalization of FP-CIT SPECT brain images using the α -stable distribution. *NeuroImage* 65, 449–455. doi: 10.1016/j.neuroimage.2012.10.005
- Saxena, P., Pavel, D. G., Quintana, J. C., and Horwitz, B. (1998). "An automatic threshold-based scaling method for enhancing the usefulness of Tc-HMPAO SPECT in the diagnosis of Alzheimer's disease," in *Medical Image Computing and Computer-Assisted Intervention — MICCAI'98*, Lecture Notes in Computer Science, eds W. M. Wells, A. Colchester, and S. Delp (Cambridge MA: Springer Berlin Heidelberg), 623–630. doi: 10.1007/BFb0056248
- Schreckenberger, M., Hägele, S., Siessmeier, T., Buchholz, H.-G., Armbrust-Henrich, H., Rösch, F., et al. (2004). The dopamine D2 receptor ligand 18F-desmethoxyfallypride: an appropriate fluorinated PET tracer for the differential diagnosis of parkinsonism. *Eur. J. Nucl. Med. Mol. Imaging* 31, 1128–1135. doi: 10.1007/s00259-004-1465-5
- Schrouff, J., Rosa, M. J., Rondina, J. M., Marquand, A. F., Chu, C., Ashburner, J., et al. (2013). PRoNTTo: pattern recognition for neuroimaging toolbox. *Neuroinformatics* 11, 319–337. doi: 10.1007/s12021-013-9178-1
- Segovia, F., Górriz, J. M., Ramírez, J., Álvarez, I., Jiménez-Hoyuela, J. M., and Ortega, S. J. (2012). Improved Parkinsonism diagnosis using a partial least squares based approach. *Med. Phys.* 39, 4395–4403. doi: 10.1118/1.4730289
- Segovia, F., Górriz, J. M., Ramírez, J., Martínez-Murcia, F. J., Levin, J., Schuberth, M., et al. (2017). Multivariate analysis of 18F-DMFP PET data to assist the diagnosis of parkinsonism. *Front. Neuroinformatics* 11:23. doi: 10.3389/fninf.2017.00023
- Segovia, F., Górriz, J. M., Ramírez, J., Salas-González, D., Álvarez, I., López, M., et al. (2010). Classification of functional brain images using a GMM-based multi-variate approach. *Neurosci. Lett.* 474, 58–62. doi: 10.1016/j.neulet.2010.03.010

- Segovia, F., Illán, I. A., Górriz, J. M., Ramírez, J., Rominger, A., and Levin, J. (2015). Distinguishing Parkinson's disease from atypical parkinsonian syndromes using PET data and a computer system based on support vector machines and Bayesian networks. *Front. Comput. Neurosci.* 9:137. doi: 10.3389/fncom.2015.00137
- Spiegel, J., Hellwig, D., Samnick, S., Jost, W., Möllers, M.-O., Fassbender, K., et al. (2007). Striatal FP-CIT uptake differs in the subtypes of early Parkinson's disease. *J. Neural Transm.* 114, 331–335. doi: 10.1007/s00702-006-0518-2
- Towey, D. J., Bain, P. G., and Nijran, K. S. (2011). Automatic classification of 123I-FP-CIT (DaTSCAN) SPECT images. *Nucl. Med. Commun.* 32, 699–707. doi: 10.1097/MNM.0b013e328347cd09
- Vapnik, V. N. (2000). *The Nature of Statistical Learning Theory*. New York, NY: Springer.
- Varma, S., and Simon, R. (2006). Bias in error estimation when using cross-validation for model selection. *BMC Bioinformatics* 7:91. doi: 10.1186/1471-2105-7-91
- Winogrodzka, A., Bergmans, P., Booij, J., van Royen, E. A., Janssen, A. G., and Wolters, E. C. (2001). [123I]FP-CIT SPECT is a useful method to monitor the rate of dopaminergic degeneration in early-stage Parkinson's disease. *J. Neural Transm.* 108, 1011–1019. doi: 10.1007/s007020170019
- Zhang, Y., Brady, M., and Smith, S. (2001). Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. *IEEE Trans. Med. Imaging* 20, 45–57. doi: 10.1109/42.906424

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.

Copyright © 2017 Segovia, Górriz, Ramírez, Martínez-Murcia and Salas-Gonzalez. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Random Forest Algorithm for the Classification of Neuroimaging Data in Alzheimer's Disease: A Systematic Review

Alessia Sarica^{1*}, Antonio Cerasa¹ and Aldo Quattrone^{1,2}

¹ Institute of Bioimaging and Molecular Physiology, National Research Council, Catanzaro, Italy; ² Institute of Neurology, University Magna Graecia, Catanzaro, Italy

OPEN ACCESS

Edited by:

Juan Manuel Gorriz,
University of Granada, Spain

Reviewed by:

Stavros I. Dimitriadis,
Institute of Psychological Medicine
and Clinical Neurosciences, Cardiff
University, United Kingdom

Feng Liu,
Tianjin Medical University General
Hospital, China

*Correspondence:

Alessia Sarica
sarica@unicz.it

Received: 23 June 2017

Accepted: 22 September 2017

Published: 06 October 2017

Citation:

Sarica A, Cerasa A and Quattrone A
(2017) Random Forest Algorithm for
the Classification of Neuroimaging
Data in Alzheimer's Disease: A
Systematic Review.
Front. Aging Neurosci. 9:329.
doi: 10.3389/fnagi.2017.00329

Objective: Machine learning classification has been the most important computational development in the last years to satisfy the primary need of clinicians for automatic early diagnosis and prognosis. Nowadays, Random Forest (RF) algorithm has been successfully applied for reducing high dimensional and multi-source data in many scientific realms. Our aim was to explore the state of the art of the application of RF on single and multi-modal neuroimaging data for the prediction of Alzheimer's disease.

Methods: A systematic review following PRISMA guidelines was conducted on this field of study. In particular, we constructed an advanced query using boolean operators as follows: ("random forest" OR "random forests") AND neuroimaging AND ("alzheimer's disease" OR alzheimer's OR alzheimer) AND (prediction OR classification). The query was then searched in four well-known scientific databases: Pubmed, Scopus, Google Scholar and Web of Science.

Results: Twelve articles—published between the 2007 and 2017—have been included in this systematic review after a quantitative and qualitative selection. The lesson learnt from these works suggest that when RF was applied on multi-modal data for prediction of Alzheimer's disease (AD) conversion from the Mild Cognitive Impairment (MCI), it produces one of the best accuracies to date. Moreover, the RF has important advantages in terms of robustness to overfitting, ability to handle highly non-linear data, stability in the presence of outliers and opportunity for efficient parallel processing mainly when applied on multi-modality neuroimaging data, such as, MRI morphometric, diffusion tensor imaging, and PET images.

Conclusions: We discussed the strengths of RF, considering also possible limitations and by encouraging further studies on the comparisons of this algorithm with other commonly used classification approaches, particularly in the early prediction of the progression from MCI to AD.

Keywords: random forest, Alzheimer's disease, mild cognitive impairment, neuroimaging, classification

INTRODUCTION

The Alzheimer's disease (AD), a common form of dementia, is a progressive neurodegenerative disorder that affects mostly elderly people (Berchtold and Cotman, 1998). It is characterized by a decline in cognitive function, including progressive loss of memory, reasoning, and language (Collie and Maruff, 2000). Mild cognitive impairment (MCI) is an intermediate state between healthy aging and AD, which is not severe enough to interfere with daily life. Although not all MCI subjects develop to AD and they remain cognitively stable for many years, the incidence of progression is evaluated between 10 and 15% per year (Palmqvist et al., 2012). There is no generally accepted cure for AD, but several treatments exist for delaying its course. For this reason, it is extremely important to early detect the MCI subjects that are at imminent risk of conversion to AD.

The diagnosis of AD is based primarily on multiple variables and factors, such as, demographics and genetic information, neuropsychological tests, cerebrospinal fluid (CSF) biomarkers, and brain imaging data. Moreover, for the assessment of the risk of conversion from MCI, the rate of change of these variables could represent a further source of knowledge. In particular, the neuroimaging technologies, such as, magnetic resonance imaging (MRI), functional MRI (fMRI), diffusion tensor imaging (DTI), single photon emission tomography (SPECT), and positron emission tomography (PET) have been widely and successfully applied in the study of MCI and AD (Greicius et al., 2004; Matsuda, 2007; Fripp et al., 2008; Frisoni et al., 2010; Acosta-Cabronero and Nestor, 2014). The choice of the neuroimaging modality depends on the duration and severity of the disease, for example when MRI could not reveal any brain alterations, fMRI, SPECT, or PET are able to assess metabolic abnormalities and DTI could be used for investigating the microstructural disruption of the white matter (WM).

The high dimension of all the features considered in the diagnosis of AD and in the progression from MCI, and their complex interactions make it very difficult for humans to interpret the data. Computer aided diagnosis (CAD) represents a valuable automatic tool for supporting the clinicians by teaching to computers to predict incipient AD. Machine learning and pattern recognition algorithms have been proven to efficiently classify AD patients and healthy controls (HC) and to distinguish between stable MCI (sMCI) subjects and progressive MCI (pMCI) that converted to AD (Zhang et al., 2012; Falahati et al., 2014; Trzepacz et al., 2014). In general, the machine learning methods used on neuroimaging data rely on a single classifier, such as, the widely used Support Vector Machine (SVM), Linear Discriminant Analysis (LDA), or Naïve Bayes. However, in the last years, ensembles algorithms resulted to be a reliable alternative to single classifiers showing better performance than the latter, especially when multi-modality variables are combined together. Although among all ensembles approaches Random Forest (RF) (Breiman, 2001) produced the best accuracies in many scientific fields (Menze et al., 2009; Calle et al., 2011; Chen et al., 2011) and in other neurological diseases (Sarica et al., 2017), it is still poorly applied in the prediction of AD, and only lately researchers paid their attention to it. In particular, RF showed

important advantages over other methodologies regarding the ability to handle highly non-linearly correlated data, robustness to noise, tuning simplicity, and opportunity for efficient parallel processing (Caruana and Niculescu-Mizil, 2006). Moreover, RF presents another important characteristic: an intrinsic feature selection step, applied prior to the classification task, to reduce the variables space by giving an importance value to each feature.

For all these reasons, the main goal of this systematic review was to highlight the role of RF as the ideal candidate for handling the high-dimensional problem and the variable redundancy in the early diagnosis of AD. We sought to review the literature in this area to identify all the works that applied the RF algorithm on single and multi-modality neuroimaging data, eventually combined with demographics and genetic information, and with neuropsychological scores. Our aim was also to evaluate how well, in term of accuracy, RF was able to classify AD and to distinguish between sMCI and pMCI, and how its intrinsic feature selection procedure could improve this overall accuracy.

Random Forest Algorithm

RF (see **Figure 1** for an illustration) is a collection or ensemble of Classification and Regression Trees (CART) (Breiman et al., 1984) trained on datasets of the same size as training set, called *bootstraps*, created from a random resampling on the training set itself. Once a tree is constructed, a set of bootstraps, which do not include any particular record from the original dataset [*out-of-bag* (OOB) samples], is used as test set. The error rate of the classification of all the test sets is the OOB estimate of

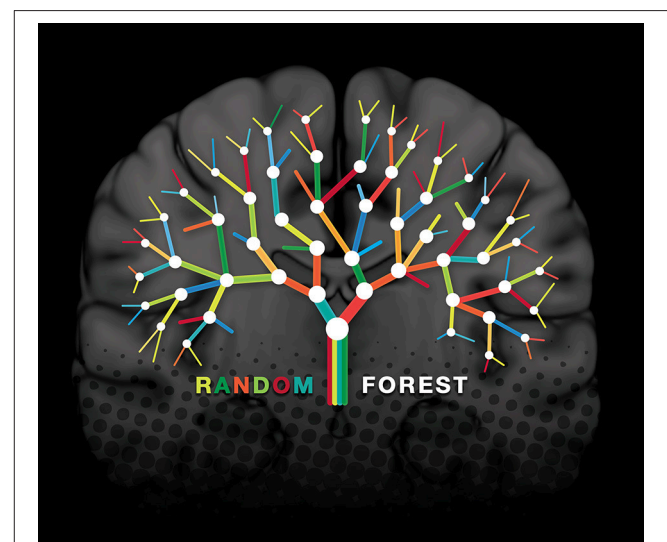
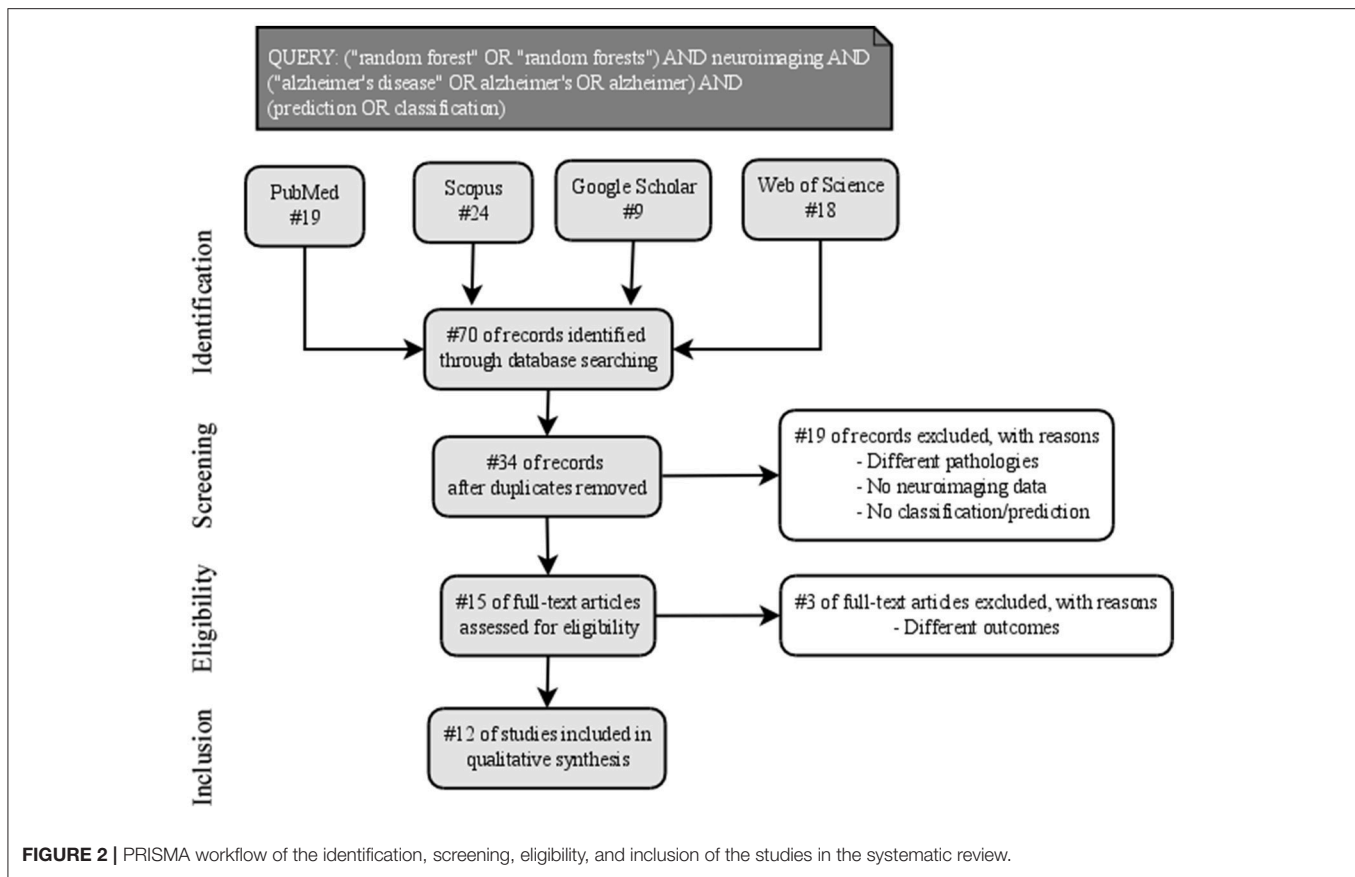


FIGURE 1 | Illustration of a random forest construct superimposed on a coronal slice of the MNI 152 (Montreal Neurological Institute) standard template. Each binary node (white circles) is partitioned based on a single feature, and each branch ends in a terminal node, where the prediction of the class is provided. The different colors of the branches represent each of the trees in the forest. The final prediction for a test set is obtained by combining with a majority vote the predictions of all single trees.



the generalization error. Breiman (1996) showed by empirical evidence that, for the bagged classifiers, the OOB error is accurate as using a test set of the same size as the training set. Thus, using the OOB estimate removes the need for a separate test set. To classify new input data, each individual CART tree (colored branches in **Figure 1**) votes for one class and the forest predicts the class that obtains the plurality of votes.

RF follows specific rules for tree growing, tree combination, self-testing and post-processing, it is robust to overfitting and it is considered more stable in the presence of outliers and in very high dimensional parameter spaces than other machine learning algorithms (Caruana and Niculescu-Mizil, 2006; Menze et al., 2009). The concept of variable importance is an implicit feature selection performed by RF with a random subspace methodology, and it is assessed by the Gini impurity criterion index (Ceriani and Verme, 2012). The Gini index is a measure of prediction power of variables in regression or classification, based on the principle of impurity reduction (Strobl et al., 2007); it is non-parametric and therefore does not rely on data belonging to a particular type of distribution. For a binary split (white circles in **Figure 1**), the Gini index of a node n is calculated as follows:

$$\text{Gini}(n) = 1 - \sum_{j=1}^2 (p_j)^2$$

where p_j is the relative frequency of class j in the node n .

For splitting a binary node in the best way, the improvement in the Gini index should be maximized. In other words, a low Gini (i.e., a greater decrease in Gini) means that a particular predictor feature plays a greater role in partitioning the data into the two classes. Thus, the Gini index can be used to rank the importance of features for a classification problem.

METHODS

For the present systematic review, we followed the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) guidelines (Liberati et al., 2009; Moher et al., 2009). The statement consists of a checklist of recommended items to be reported and a four-step flow diagram (**Figure 2**).

Published titles and abstracts in the English language from the first of January 2007 to the first of May 2017 were searched systematically across the following databases: PubMed, Scopus, Google Scholar, and Web of Science. The search terms were concatenated in an advanced query using boolean operators as follows: ("random forest" OR "random forests") AND neuroimaging AND ("alzheimer's disease" OR alzheimer's OR alzheimer) AND (prediction OR classification). After the initial web search, duplicate items among databases were removed.

During the screening phase, to be assessed for eligibility, studies were required to: (1) investigate a cohort of AD in cross-sectional case-control or longitudinal design, (2) analyze

neuroimaging data, (3) apply RF algorithm as Machine Learning technique for the classification of AD patients.

To reduce a risk of bias, two authors (A.S. and A.C.) independently screened paper abstracts and titles, and analyzed the full papers that met the inclusion criteria, as suggested by the PRISMA guidelines. The reference lists of examined full-text papers were also scrutinized for additional relevant publications.

Data extracted from the studies—finally included in the qualitative synthesis—were: (1) sample diagnosis, (2) sample size and mean age, (3) neuroimaging acquisition type, (4) features of interest, (5) RF classification parameters, (6) classification performance validation, and (7) selected findings in terms of classification performance.

RESULTS

Study Selection

Figure 2 reported the four phases—identification, screening, eligibility and inclusion—of the process for the selection of the studies in this review. Nineteen records were excluded after the initial screening of title and abstract and three more records were removed after the full-text assessment, following the inclusion criteria. Finally, 12 studies were included in qualitative synthesis.

Study Characteristics

Data extracted from the studies were summarized in **Table 1**. In particular, we reported those characteristics that are related to the highest performance reached by RF in each study. Regarding the cohort diagnosis, two works (Tripoliti et al., 2007; Lebedev et al., 2014) investigated Alzheimer's patients (AD) and healthy controls (HC), four works (Cabral et al., 2013; Sivapriya et al., 2015; Maggipinto et al., 2017; Son et al., 2017) had AD, HC, and MCI, two studies (Gray et al., 2013; Moradi et al., 2015) considered AD, HC, stable MCI (sMCI), and progressive MCI (pMCI, converted to AD), two had sMCI and pMCI (Wang et al., 2016; Ardekani et al., 2017), one had HC and MCI (Lebedeva et al., 2017) and one (Oppedal et al., 2015) had AD, HC, and Lewy-body dementia (LBD) patients.

All studies, except two (Cabral et al., 2013; Maggipinto et al., 2017), which used FDG-PET and DTI acquisition respectively, investigated structural MRI data alone (Lebedev et al., 2014; Moradi et al., 2015; Ardekani et al., 2017; Lebedeva et al., 2017) or in combination with features extracted from other modalities, that is FDG-PET (Gray et al., 2013; Sivapriya et al., 2015), florbetapir-PET (Wang et al., 2016), FLAIR (Oppedal et al., 2015) and fMRI (Tripoliti et al., 2007; Son et al., 2017).

Eight works (Tripoliti et al., 2007; Cabral et al., 2013; Lebedev et al., 2014; Moradi et al., 2015; Sivapriya et al., 2015; Ardekani et al., 2017; Lebedeva et al., 2017; Maggipinto et al., 2017) applied feature selection/elimination for reducing the dimension of the variables space. The number of trees used in the RF was not specified in two cases (Moradi et al., 2015; Son et al., 2017). Finally, we reported in the column Results of **Table 1** the—highest—overall accuracies of binary or ternary classifiers reached by each study, except for the one (Tripoliti et al., 2007) that provided only sensitivity and specificity. **Figure 3** presented a comparison—where applicable—of accuracies obtained by the

studies for the binary models AD vs. HC (**Figure 3A**, with a mean of 88.8%), MCI vs. HC (**Figure 3B**, with a mean of 79%), sMCI vs. pMCI (**Figure 3C**, with a mean of 74%), and the multi-class problem AD vs. HC vs. MCI (**Figure 3D**, with a mean of 71.42%)

More details about individual works, such as, the results obtained with other algorithms or other subsets of features, could be found in the next section.

Results of Individual Studies

Tripoliti et al. (2007)

Tripoliti et al. (2007) conducted a study on 41 subjects, divided into three groups: 12 subjects were AD patients (mean age 77.2, 7 females), from very mild to mild following the Clinical Dementia Rating (CDR = 0.5/1), 14 subjects were healthy young controls (mean age 21.1, 9 females, CDR = 0) and 14 were healthy elderly subjects (mean age 74.9, 9 females, CDR = 0).

The cohort underwent a visual fMRI finger tapping task. Raw structural and functional images were preprocessed for correction of motion artifacts, registered and normalized.

Demographic and behavioral data were grouped with the features extracted from the data preprocessing phase: (i) head motions parameters; (ii) volumetric measures, i.e., volumes obtained from the segmentation of gray matter (GM), WM and CSF; (iii) activation patterns, consisting in several measures derived from the activated voxels and clusters; (iv) hemodynamic measures extracted from the BOLD responses, such as, the amplitude of venous volume or of vascular signal. Authors applied a feature selection on this dataset for reducing the dimensionality by removing highly correlated variables. Selected features were used for training a RF classifier with 10 trees, and the performance was assessed using 10-fold cross-validation accuracy. Two separated datasets were evaluated: the first consisted of AD patients and both young and old healthy subjects, while the second consisted of AD and only old controls. Sensitivity and specificity of the two binary classifiers were ranging from 94 to 98%, depending of the subset of selected features. The highest values were obtained on the dataset that included AD and old controls, with a 98% of both sensitivity and specificity.

Gray et al. (2013)

Gray et al. (2013) selected a cohort of 147 subjects from the ADNI database, consisting of 37 AD patients (mean age 76.8, 14 females, CDR = 0.5/1), 75 MCI patients divided into 34 stable MCI (sMCI, mean age 75.7, 12 females, CDR = 0.5) and 41 subjects progressed to AD (pMCI, mean age 76.1, 12 females, CDR = 0.5), and 35 HC (HC, mean age 74.5, 12 females, CDR = 0). All subjects underwent morphological 1.5T MRI, FDG-PET, and CSF analysis at the baseline and authors used already pre-processed data by ADNI. In particular, structural MRI and FDG-PET images were motion-corrected, examined for major artifacts and registered to the standard space MNI. Eighty-three volumetric region-based features were extracted from MRI, while signal intensities of 239,304 voxels were obtained from FDG-PET. Biological features were CSF-derived measures of A β , tau, and ptau. Furthermore, a categorical variable describing the ApoE genotype was used as genetic feature.

TABLE 1 | Characteristics of each of the twelve studies included in the systematic review.

| Study | Cohort | Nr. of subjects (mean age, nr. of females) | Neuroimaging acquisition | Features of interest | Classification parameters | Performance validation | Results |
|---------------------------|--------------------------|--|---|---|--|--|--|
| Tripoliti et al., 2007 | AD HC | 12 (77.2, 7) 14 (74.9, 9) | – 1.5 T MRI – Task-based fMRI | – Demographic data; – Behavioral data; – Head motions parameters; – Volumetric measures; – Activation patterns; – BOLD-derived hemodynamic measures. | – Feature selection based on correlation; – RF with 10 trees. | 10-fold cross-validation sensitivity/ specificity | AD vs. HC: 98%/98% |
| Gray et al., 2013 | AD sMCI pMCI HC | 37 (76.8, 14) 34 (75.7, 12) 41 (76.1, 12) 35 (74.5, 12) | – 1.5 T MRI – FDG-PET | – Volumetric measures; – FDG-PET voxel intensities whole-brain; – CSF-derived measures; – Genetic information. | – RF with 5,000 trees. | Stratified repeated random sampling accuracy on a separate test set | AD vs. HC: 89% MCI vs. HC: 74.6% sMCI vs. pMCI: 58.4% |
| Cabral et al., 2013 | AD MCI HC | 59 (78.2, 25) 59 (77.7, 19) 59 (77.4, 21) | – FDG-PET | – FDG-PET voxel intensities; | – Feature selection with Mutual Information criterion; – Decomposition by the one-vs.-all scheme; – Aggregation scheme with voting strategy (MAX); – RF with 100 trees. | Repeated 10-fold cross-validation accuracy | AD vs. MCI vs. HC: 64.63% |
| Lebedev et al., 2014 | AD HC | 185 (75.2, 92) 225 (75.95, 110) | – 1.5 T MRI | – Non-cortical volumes; – Cortical thickness; Jacobian maps; – Sulcal depth. | – Recursive feature elimination with Gini index; – RF with 1,000 trees. | Overall accuracy on a separate test set | AD vs. HC: 90.3% |
| Moradi et al., 2015 | AD sMCI pMCI HC | 200(55-91, 97) 100 (57-89, 34) 164 (77-89, 67) 231 (59-90, 112) | – 1.5 T MRI | – GM density values; – Age; – Neuropsychological scores. | – Feature selection with regularized logistic regression framework | 10-fold cross-validation accuracy | sMCI vs. pMCI: 82% |
| Oppedal et al., 2015 | AD LBD HC | 57 (N.A.) 16 (N.A.) 36 (N.A.) | – 1.0/1.5 T MRI – FLAIR | – Local binary pattern (LBP); – Image contrast measure (C). | – RF with 10 trees. | 10-fold nested cross-validation accuracy | AD vs. LBD vs. HC: 87% AD+LBD vs. HC: 98% AD vs. LBD: 74% |
| Sivapriya et al., 2015 | AD MCI HC | 140 (N.A.) 450 (N.A.) 280 (N.A.) | – MRI – FDG-PET | – Volumetric measures; – FDG-PET uptake ROI-based; – Neuropsychological scores. | – Feature selection with particle swarm optimization approach coupled with the Merit Merge technique (CPEMM); – RF with 100 to 1,000 trees. | 5-fold cross-validation accuracy | AD vs. MCI vs. HC: 96.3% |
| Wang et al., 2016 | sMCI pMCI | 65 (72.2, 26) 64 (72.5, 29) | – 1.5 T MRI – florbetapir- PET – FDG-PET | – Morphological measures; – florbetapir-PET uptake whole-brain; – FDG-PET uptake whole-brain. | – RF with 500 trees. | Leave-one-out cross-validation accuracy | sMCI vs. pMCI: 73.64% |

(Continued)

TABLE 1 | Continued

| Study | Cohort | Nr. of subjects (mean age, nr. of females) | Neuroimaging acquisition | Features of interest | Classification parameters | Performance validation | Results |
|----------------------------|-----------------|--|-----------------------------|--|---|---|-----------------------------------|
| Ardekani et al., 2017 | sMCI pMCI | 78 (74.75, 24) 86 (74.10, 31) | – 1.5 T MRI | – Hippocampal volumetric integrity; – Neuropsychological scores. | – Feature selection with Gini index; – RF with 5,000 trees. | OOB estimation of classification accuracy | sMCI vs. pMCI: 82.3% |
| Lebedeva et al., 2017 | MCI HC | 32 (78.1, 22) 40 (76.4, 29) | – 1.5/3 T MRI | – Cortical thickness; – Subcortical volumes. – MMSE | – Feature selection with Gini index; – RF with 5,000 trees. | OOB estimation of classification accuracy | MCI vs. HC: 81.3% |
| Maggipinto et al., 2017 | AD MCI HC | 50 (N.A.) 50 (N.A.) 50 (N.A.) | – DTI | – TBSS FA voxels; | – Feature selection with the Wilcoxon rank sum test and Relieff algorithm; – RF with 300 trees. | Repeated 5-fold cross-validation accuracy | AD vs. HC: 87% MCI vs. HC: 81% |
| Son et al., 2017 | AD MCI HC | 30 (74, 18) 40 (74.3, 21) 35 (76.06, 23) | – 3 T MRI – rs-fMRI | – Subcortical volumes; – Eigenvector centrality of functional networks ROI-based. | N.A. | Repeated leave-one-out cross-validation accuracy | AD vs. MCI vs. HC: 53.33% |

Data are related to the highest performance reached by random forest. AD, Alzheimer's disease; HC, healthy controls; MCI, Mild cognitive impairment; cMCI, converter MCI; pMCI, progressive MCI; LBD, Lewy-body dementia; MRI, Magnetic resonance imaging; fMRI, functional MRI; rs-fMRI, resting state fMRI; PET, positron emission tomography; FDT-PET, fluorodeoxyglucose PET; DTI, Diffusion tensor imaging; GM, Gray matter; ROI, Region of interest; MMSE, Mini mental state examination; TBSS, Tract-based spatial statistics; OOB, out-of-bag; N.A., not applicable.

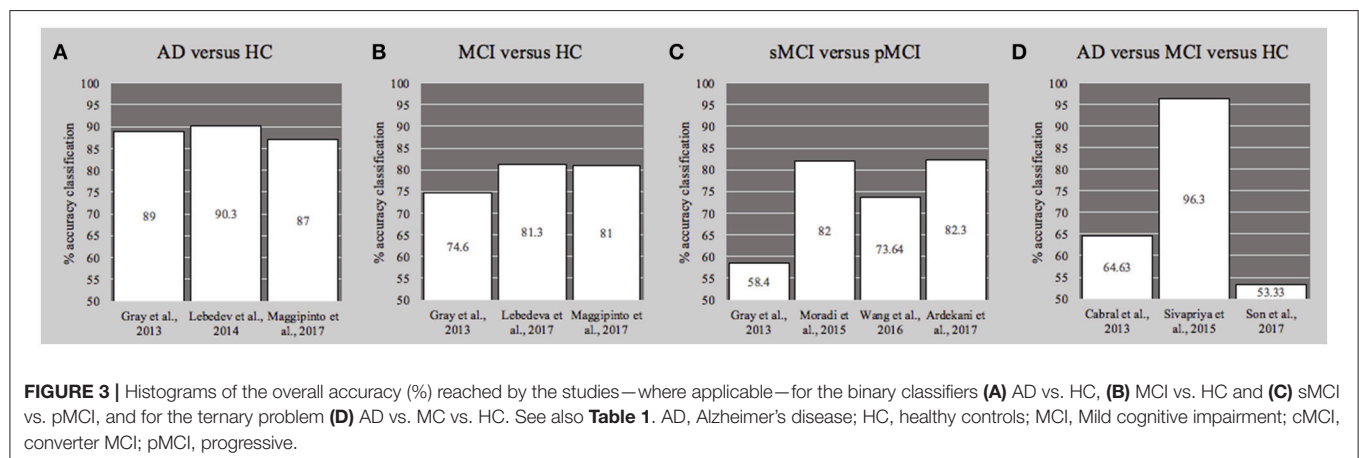


FIGURE 3 | Histograms of the overall accuracy (%) reached by the studies—where applicable—for the binary classifiers **(A)** AD vs. HC, **(B)** MCI vs. HC and **(C)** sMCI vs. pMCI, and for the ternary problem **(D)** AD vs. MC vs. HC. See also **Table 1**. AD, Alzheimer's disease; HC, healthy controls; MCI, Mild cognitive impairment; cMCI, converter MCI; pMCI, progressive.

Three different binary datasets were used for the RF classification: AD vs. HC, MCI vs. HC, sMCI vs. pMCI. The performance of each classifiers was evaluated with a stratified repeated random sampling approach, where, in each of the 100 runs, the dataset was divided into training (75%) and test set (25%). Accuracy on the test set was then calculated as the mean of all the 100 repetitions. The RF models were trained with 5,000 trees on the feature data from each of the four modalities independently and the feature importance ranking was extracted. As further analysis, authors measured the similarity between pairs of examples from the RF classifiers and applied a Manifold learning approach on data from single-modality and on combined/concatenated features (multi-modality).

Although the single-modality classification results were comparable between the original dataset and the embedded feature one, the latter presented the best performances as following: 86.4% for the AD vs. HC with the FDG-PET data, 73.8% for MCI vs. HC with the genetic data, 58.4% for the sMCI vs. pMCI with MRI data. A slight increase of the accuracy was obtained with the multi-modality classification for AD vs. HC (+2.6%) and for MCI vs. HC (+0.8%), while for pMCI vs. sMCI there was a small decrease (−0.4%).

Cabral et al. (2013)

Cabral et al. (2013) collected 177 subjects from the ADNI database, divided into three balanced groups: AD patients (mean age 78.2, 25 females, CDR > 0.5), MCI patients (MCI,

mean age 77.7, 19 females, CDR = 0.5) and HC (mean age 77.4, 21 females, CDR = 0). Authors analyzed FDG-PET data, acquired 24 months after the first visit and already pre-processed by ADNI. In particular, they used the voxel intensities (VI) as features of interest, for a total of 309,881 variables. The original dataset was decomposed by using the one-vs.-all scheme, resulting into three subsets: AD vs. ALL, MCI vs. ALL, HC vs. ALL. The Mutual Information criterion was used for extracting the optimal features with the highest ranking value, separately for each pairwise problem. The selected features were then used for training three binary RF models with 100 trees. As aggregation scheme for the ternary problem, the voting strategy (MAX) was applied. The classification performance was then assessed by the 10-fold cross-validation accuracy, repeated 5 times with fold randomization. The ternary RF classifier provided a multiclass accuracy of 64.63%. It must be addressed that the authors applied other two algorithms, linear and RBF SVM, obtaining, respectively, an accuracy of 66.33% and 66.78%.

Lebedev et al. (2014)

The study of Lebedev et al. (2014) was based on a cohort of 575 subjects from ADNI database, divided into three main groups: 185 AD (mean age 75.2, 92 females, CDR = 1), 165 patients with MCI (mean age 75.46, 62 females, CDR = 0.5) of which 149 progressed to AD within 4 years, and 225 HC (mean age 75.95, 110 females, CDR = 0). The MCI group was split into six subgroups according to the month of MCI-to-AD conversion (6th-, 12th-, 18th-, 24th-, 36+th-month converters and non-converters).

The features of interest were extracted from 1.5 T MRI images using a surface-based cortex reconstruction and volumetric segmentation. In particular, (i) non-cortical volumes, (ii) cortical thickness (CTH), (iii) Jacobian maps and (iv) sulcal depth were measured for each subject. The ability of these parameters in distinguishing AD from HC, was assessed individually and with a combination of measurements of CTH and non-cortical volumes.

The feature importance was assessed with the intrinsic characteristic of RF consisting of the recursive feature elimination (RFE) with the Gini index as criterion and 10,000 trees. The performance of models—with and without RFE—was evaluated as the overall accuracy on a separate test set with 35 AD and 75 HC. Findings revealed that the highest accuracy (90.3%) for the classifier AD vs. HC was obtained with the RFE on the combined dataset with thickness and non-cortical volumes. An increase of 0.7% was found in this accuracy when authors combined all models by a majority vote approach. The majority vote method resulted to have also the best ability to predict MCI-to-AD conversion 2 years before actual dementia onset with sensitivity/specificity of 76.6/75%. As further analysis, authors found that the adding of ApoE genotype and demographics data did not improve the overall accuracy in distinguishing AD from HC, while it showed an increase of sensitivity/specificity (83.3/81.3%) in the prediction of MCI conversion.

Moradi et al. (2015)

Moradi et al. (2015) obtained baseline data for their analysis from the ADNI database and they selected 825 subjects grouped as: 200 AD patients (age range 55–91, 97 females), 100 stable MCI (sMCI, age range 57–89, 34 females), 164 MCI progressed to AD within 3 years from the baseline (pMCI, age range 77–89, 67 females) and 231 HC (age range 59–90, 112 females). Another group of 100 unknown MCI (uMCI, age range 54–90, 81 females) diagnosed as MCI at the baseline but with missing diagnosis at 36 months follow-up was also considered. For integrating the unlabeled group of uMCI into the training set and assigning them to the pMCI or sMCI class, the authors used a low density separation (LDS) approach for semi-supervised learning.

All subjects underwent 1.5 T MRI acquisition and the T1-w scans were preprocessed following the voxel-based morphometry approach. In particular, T1-w images were corrected, spatially normalized and segmented into GM, WM, and CSF. The GM maps were then further processed for extracting 29,852 GM density values—for each subject—used as MRI features for the classification task.

The high number of GM voxels was reduced with a feature selection approach consisted in the regularized logistic regression framework applied only on the dataset with AD and HC subjects. The selected variables were then aggregated with age and cognitive measurements and used for building the RF classifier for predicting AD in MCI patients, i.e., sMCI vs. pMCI.

The RF model performance was evaluated as the mean accuracy calculated by 10-fold cross-validation. The highest accuracy in distinguishing the MCI-to-AD conversion reached almost the 82% when the concatenated measures—age, cognitive, and voxel—and the combination of LDS and RF were considered. The importance analysis of MRI features, age, and cognitive measurements calculated by RF classifier revealed that the first three most predictive variables were: MRI voxels, the Rey's Auditory Verbal Learning Test (RAVLT) and the Alzheimer's Disease Assessment Scale—cognitive subtest 11 (ADAS-cog total-11).

Oppedal et al. (2015)

In Oppedal et al. (2015), a total of 73 mild dementia subjects, divided into 57 AD patients and 16 LBD patients, together with 36 HC were investigated. The cohort MRIs were acquired in different research centers with 1.0/1.5 T scanners and FLAIR images were also obtained. T1-w images were corrected, registered and segmented for extracting the white matter (WM) tissue. From the pre-processed FLAIR images, the WM lesions (WML) maps were automatically created. In a second phase of the study, authors applied the local binary pattern (LBP) approach as a texture descriptor on both T1 and FLAIR images and their derived WM and WML maps as ROIs. For enhancing the discriminative power of LBP, an image contrast measure (C) was added as variable for every voxel in the specified ROI. The total number of features for each subject was 48, resulting from the combination of LBP and C values in each ROI.

Feature selection and classification were performed with a RF classifier with 10 trees and the 10-fold nested cross validation accuracy was used as the performance metric. In particular, three

RF models were built: (i) a ternary problem HC vs. AD vs. LBD, (ii) a binary classifier HC vs. AD+LBD and (iii) another binary model AD vs. LBD.

For the ternary problem—HC vs. AD vs. LBD—the best accuracy (87%) was reached when the classifier was trained on the texture features extracted from the T1 images in the WML masks (T1WML). Results of the model HC vs. AD+LBD revealed that the highest accuracy (98%) was obtained also when only T1WML variables were considered. On the contrary, for distinguishing AD from LBD with the maximum accuracy (74%) the texture features should be extracted from the T1 in the WM ROI.

Sivapriya et al. (2015)

Four datasets from the ADNI database were used by Sivapriya et al. (2015) and three different groups of subjects were selected: AD, MCI, and HC. The number of subjects in each dataset varied according to the features considered: (i) Neuropsychological dataset (150 AD, 400 MCI, 200 HC), (ii) Neuroimaging dataset (250 AD, 200 MCI, 250 HC), (iii) Baseline combined data with both neuropsychological and neuroimaging measures (140 AD, 450 MCI, 280 HC), and (iv) combined dataset (150 AD, 400 MCI, 200 HC). Some of the neuropsychological tests used were the Clinical dementia ratio-SB, the ADAS, the RAVLT, and the MOCA. Authors used already pre-processed MRI data by ADNI for their study, in particular neuroimaging measures extracted from T1-w and FDG-PET images, consisting in volumes and average PIB SUVR of several regions of interest (ROIs).

The feature selection and classification task was composed by three main phases in which RF performance was evaluated together with other ensemble algorithms—Naïve Bayes, J48 and SVM. Each classifier was trained with each of the four datasets, after that they were dimensionally reduced with a particle swarm optimization approach coupled with the Merit Merge technique (CPEMM). The performance of the classification models was evaluated with the 5-fold cross-validation accuracy of the ternary problem AD vs. MCI vs. HC. RF—implemented with 100 to 1,000 trees—showed its best multi-class accuracy (96.3%) when it was trained on the baseline combined dataset and the same result was obtained with the CPEMM feature selection methodology. It must be addressed that RF reached comparable performance of the other classification algorithms, except for SVM that presented the lowest accuracies in the delineation of dementia.

Wang et al. (2016)

The study of Wang et al. (2016) included 129 subject with MCI (CDR = 0.5) from the ADNI database. The cohort was divided into 65 stable MCI (sMCI, mean age 72.2, 26 females) and 64 progressive MCI (pMCI, mean age 72.5, 29 female), who converted to AD within 3 years from the baseline. All subject underwent the acquisition of 1.5 T MRI, florbetapir-PET and FDG-PET. Authors analyzed already pre-processed neuroimaging data by ADNI, separately grouped according to the modality of acquisition, i.e., features extracted from the T1-w images and the uptake of florbetapir and FDG. A dataset with a combination of these multimodal measures was also evaluated. Three classification algorithms—partial least square

(PLS, informed and agnostic), linear SVM and RF (500 trees)—were trained on these four different datasets. Their ability in distinguishing sMCI from pMCI was assessed with the leave-one-out cross-validation accuracy. RF showed the best accuracy (73.64%) when it was trained on the combined multi-modal features dataset. A comparable result (76.74%) on the same dataset was reached by SVM. On the contrary, informed PLS generally outperformed both RF and SVM especially when the three neuroimaging modalities are fused (81.4% of accuracy).

Ardekani et al. (2017)

Ardekani et al. (2017) applied their classification task on a cohort of 164 MCI (CDR = 0.5) patients from the ADNI database, divided into 78 stable MCI (sMCI, mean age 74.75, 24 females) and 86 MCI converted to AD within 3 years from the baseline (pMCI, mean age 74.10, 31 females). All selected subjects underwent two 1.5 T MRI acquisitions, at the baseline and at ~1 year later. Neuropsychiatric scores of these two time points were also considered in the analysis. T1-w images—without any pre-processing—were used for calculating the hippocampal volumetric integrity (HVI), defined as the fraction of volume of a region that is expected to surround the hippocampus in a normal brain that is occupied by tissue (rather than CSF). The HVI is measured—separately for each hemisphere—as the area under the histogram curve for voxel values above a CSF intensity threshold. The HVI measures and the neuropsychiatric scores were merged for a total of 16 features for each subject, including their average rate of change between the baseline and the 1-year follow-up.

Several RF models (5,000 trees) were trained on different feature subsets and their performance were evaluated with the OOB estimation of classification accuracy. The mean reduction of Gini impurity index was used for the assessment of the variable importance.

The highest accuracy (82.3%) in distinguishing between sMCI and pMCI was reached when the combination of neuroimaging and neuropsychiatric features was considered as training set. The classifiers built only on the baseline measures or only on HVI values showed indeed poor performance. The variable ranking of the 16 features revealed that—according to the impurity criterion—ADAS cognitive test was the most important one, followed by the rate of change of the right HVI.

Lebedeva et al. (2017)

The work of Lebedeva et al. (2017), was aimed at predicting MCI and dementia in late-life depression (LLD) patients 1 year prior to the diagnosis. The analysis was conducted on a cohort of 32 patients (MCI-DEM, mean age 78.1, 22 females) including 21 MCI and 8 AD, and a group of 40 age—sex—matched HC (mean age 76.4, 29 females) from the PRODE prospective multicenter study (Borza et al., 2015). All subjects underwent 1.5/3 T MRI acquisition at the baseline and after 1 year. T1-w images were pre-processed for extracting CTH and subcortical volumes (SV) with a standard pipeline, for a total of 148 features. Clinical and neuropsychological assessment was performed for each subject at both time points.

Several RF models (5,000 trees) were built for classifying MCI-DEM or MCI vs. HC at 1-year follow-up, by varying the feature space, i.e., separated CTH and SV variables, a combination of CTH and SV, and with/without the addition of demographic and clinical data. The OOB overall accuracy was assessed as performance metric. The model for discriminating MCI-DEM from HC reached the best result (81.3% of OOB overall accuracy) when the CTH, SV, and MMSE values were combined together. The accuracy resulted to be higher (90.1%) in the model of MCI (excluding AD patients) vs. HC with SV and MMSE as training features. The variable importance ranking—measured with the Gini criterion—showed that, in every RF models, the most relevant features were the right ventral diencephalon, the middle anterior corpus callosum and the right hippocampus.

As further analysis, authors used their PRODE cohort (MCI-DEM and HC) as test set for the RF model previously built by Lebedev et al. (2014) on AD and HC from ADNI database. The accuracy was better (67%) when only SV measures were used than when SV and CTH were combined (57.5%).

Maggipinto et al. (2017)

The cohort investigated by Maggipinto et al. (2017) was obtained from ADNI database and it consisted of 150 subjects divided into three groups: 50 AD, 50 MCI, and 50 HC with an age range from 55 to 90. Diffusion-weighted scans acquired with a 3 T scanner was used for this machine learning study, randomly selected from the baseline and follow-up visit. DTIs were pre-processed for correction of movement artifacts and eddy currents with a standard pipeline. A diffusion tensor was fitted for each subject and fractional anisotropy (FA) and mean diffusion (MD) maps were extracted. The FA and MD maps were then used as input for a tract-based spatial statistics (TBSS) analysis, which—for each subject—produced ~120,000 voxels for each diffusion metric.

In a first phase, authors assessed the importance value of the voxels in discriminating AD from HC with two different feature selection methods: the Wilcoxon rank sum test and the ReliefF algorithm, which were used both within a non-nested and nested approach. For the classification task, fifteen subsets were then created by selecting an increasing number—from 50 to 3,000—of most discriminating voxels, ordered by decreasing importance. RF models were trained with 300 trees on each of these feature subspaces and their performance was evaluated with a repeated (100 runs) 5-fold cross-validation accuracy.

The models built on the FA features selected with the non-nested approach showed the highest accuracies in both binary problems, AD vs. HC (87%) and MCI vs. HC (81%). The non-nested variable selection resulted to produce better results than the nested one also when MD voxels were used for training the classifiers (83% for AD vs. MCI and 79% for MCI vs. HC).

Son et al (2017)

A sample of 105 subjects was selected by Son et al. (2017) from the ADNI database. The cohort was divided into three age—sex—matched groups: 30 AD (mean age 74, 18 females), 40 MCI (mean age 74.3, 21 females) and 35 HC (mean age 76.06, 23 females). All participants underwent 3 T acquisition of T1-w images and resting state functional MRI (rs-fMRI). Structural

scans were pre-processed for correcting movement artifacts and smoothed, and then they were segmented into WM, GM, and CSF. The volumes of 10 subcortical regions were calculated as measure of atrophy. The rs-fMRI images were pre-processed and registered onto the T1-w and aligned to the MNI standard space. Given a set of ROIs from the AAL atlas as nodes, the functional networks were constructed by defying the edges as correlation values between nodes. Authors quantified the connectivity of the functional networks within the 10 subcortical regions with the eigenvector centrality measure among AD and HC, MCI and HC, and AD and MCI.

The ternary problem, AD vs. MCI vs. HC, was evaluated by training a RF classifier with the SV and the eigenvector centrality measures as features. The multi-class accuracy of the RF model was assessed with a repeated (105 runs) leave-one-out cross-validation approach. Authors reached a poor performance (accuracy: 53.33%) in distinguishing among AD, MCI, and HC subjects. However, they identified distinctive regional atrophy and functional connectivity patterns characterizing each binary problem AD vs. HC (thalamus, putamen and hippocampus bilaterally and left amygdala), MCI vs. HC (left putamen and right hippocampus), and MCI vs. AD (bilateral hippocampus and right amygdala).

DISCUSSION

RF has been successfully applied in many scientific realms such as, the bioinformatics, proteomics, and genetics (Menze et al., 2009; Calle et al., 2011; Chen et al., 2011), but it was less applied on neuroimaging data for the prediction of the Alzheimer's disease. The present paper is the first, to our knowledge, that systematically analyzed the literature of the last 10 years on the use of the RF algorithm on neuroimaging data for the early diagnosis of AD. In this review, we summarized the characteristics of twelve works (Tripoliti et al., 2007; Cabral et al., 2013; Gray et al., 2013; Lebedev et al., 2014; Moradi et al., 2015; Oppedal et al., 2015; Sivapriya et al., 2015; Wang et al., 2016; Ardekani et al., 2017; Lebedeva et al., 2017; Maggipinto et al., 2017; Son et al., 2017) by focusing our attention on performance reached by their algorithms.

A direct comparison of the results of the selected works is influenced by several factors, such as, the different sample sizes, neuroimaging modalities, and different methods for the feature selection. However, we found several points in common among papers, such as, similar performance validation approaches, as well as a general trend showing that the classification based on a combination of features extracted from different categories improved the ability in predicting AD. Another important common aspect of the selected articles is the use of data from the ADNI database. Indeed, 10 works (Cabral et al., 2013; Gray et al., 2013; Lebedev et al., 2014; Moradi et al., 2015; Sivapriya et al., 2015; Wang et al., 2016; Ardekani et al., 2017; Lebedeva et al., 2017; Maggipinto et al., 2017; Son et al., 2017) applied their methodologies on ADNI cohorts.

The best accuracies—around 90%—for the binary problem AD vs. HC, were observed when the RF classifiers were trained on high-dimensional and multi-modality data (Tripoliti et al.,

2007; Cabral et al., 2013; Gray et al., 2013; Lebedev et al., 2014). Superior performance of these models can be explained by the ability of RF to detect less extensive changes in the variables, which could be not revealed by others algorithms. Moreover, Moradi et al. (2015) showed that RF was more immune to the data type thanks to its capability to handle discrete data and to apply an efficient discretization algorithm on continuous data type before the learning step.

The binary models for distinguishing MCI from HC and stable MCI from progressive MCI showed lower accuracies, around 82%, although it was similarly improved by multi-modal data classification (**Figure 3**). In particular, the inclusion of age as well as cognitive measurements (MMSE and ADAS-cog), in the space of features, significantly increased the classification of MCI vs. HC (Gray et al., 2013; Lebedeva et al., 2017; Maggipinto et al., 2017) and the AD conversion prediction in MCI patients (Moradi et al., 2015; Wang et al., 2016; Ardekani et al., 2017). On the contrary, for the conundrum between sMCI vs. pMCI, Gray et al. (2013) found that the accuracy reached on multi-modality classification is not significantly different from that obtained with MRI information alone. Interestingly, authors suggested that the lack of improvement in distinguishing the progression to AD, could be overcome by incorporating longitudinal information, as indeed Ardekani et al. (2017) demonstrated afterwards by considering the rate of change of variables.

Three works (Cabral et al., 2013; Sivapriya et al., 2015; Son et al., 2017) investigated the ternary problem: AD vs. MCI vs. HC, but only the work of Sivapriya et al. (2015) reached a reliable accuracy of 96.3%. The low performance of the other two studies—64.63% of Cabral et al. (2013) and 53.33% of Son et al. (2017)—might be due to the heterogeneous pattern of brain changes across the three groups and the inability of RF to model the too large variability in the stages of pathological process. Thus, although RF can be naturally extended to multi-class problems, the AD vs. MCI vs. HC ternary model could not be still translated into a real-world clinical scenario.

Another interesting observation was that, both in binary and ternary problems, feature selection based on the Gini index, improved the overall performance and this is true also for the works in which only a neuroimaging modality was used (Lebedev et al., 2014; Ardekani et al., 2017; Lebedeva et al., 2017; Maggipinto et al., 2017). Other kinds of feature selection and extraction, applied prior to the RF classification, showed also an improvement in the overall accuracies (Tripoliti et al., 2007; Cabral et al., 2013; Moradi et al., 2015; Oppedal et al., 2015; Sivapriya et al., 2015; Wang et al., 2016; Ardekani et al., 2017; Lebedeva et al., 2017; Maggipinto et al., 2017).

A further interesting characteristic of the RF algorithm in the AD realm was the estimates of the features importance. The ranking of the variables plays an important role because it could assess which of the features contribute most to the prediction by also providing a correspondence to anatomical regions or structures with a biologically plausible connection to pathology (Gray et al., 2013; Lebedev et al., 2014; Moradi et al., 2015; Ardekani et al., 2017).

A limitation of this systematic review concerns the lack of information about the tuning of the RF parameters. In particular,

poor information were reported in the selected works about how the number and depth of trees in the forest or the splitting criteria were chosen. Although, this tuning is performed automatically by RF, how external assessment of these parameters (i.e., cross-validation approach) would improve the overall accuracies is still unknown.

Again, what still remains to be assessed is the performance of RF algorithm on multi-site data. As already demonstrated for rs-fMRI datasets from different sites (Abraham et al., 2017; Dansereau et al., 2017), the accuracy and the reliability of the biomarkers extraction could be enhanced by dramatically increasing the cohort size. Moreover, it was shown that classifiers trained on data from multiple sources will likely generalize better to new observations (Dansereau et al., 2017), avoiding the overfitting. Thus, it would be interesting to evaluate how well RF could classify when it is trained on features that are not invariant across sites and how the sample heterogeneity influences its performance.

This systematic review provided, for the first time, a framework for the exploration of the RF algorithm and of its strength in predicting AD when high-dimensional and multi-modal neuroimaging data are combined with demographics, genetic and cognitive scores. Indeed, as recently stated by Rathore et al. (2017), no single neuroimaging modality is enough to reach optimal accuracy for automatic AD prediction, but only through the combination of different methodologies, the classification task could be effectively translated into the clinical realm. Our work supported the idea that there is some complementary information between modalities and that this knowledge can be successfully explored with a combination of classifiers rather than a single one. The RF, as a bagging ensemble model, provided promising results, but with possible limitations. Thus, given the high accuracies reached by RF in the classification of dementia, we aimed at encouraging further studies, especially for comparing and integrating this algorithm with other machine learning approaches, such as, the deep learning, which recently showed its potentiality in the investigation of neuroimaging correlates (Shen et al., 2017; Vieira et al., 2017). In the future, the aggregation of multi-approaches (RF, Deep-learning and SVM), multimodal (MRI, DTI, PET) and multi-sites data would drastically increase our ability to extract reliable biomarkers of neurodegenerative diseases.

AUTHOR CONTRIBUTIONS

AS: Research project: Conception, Organization, and Execution. Statistical Analysis: Design, Execution, Review, and Critique. Manuscript: Writing of the first draft, Review, and Critique. AC: Research project: Conception, Organization and Execution. Manuscript: Review and Critique. AQ: Research project: Organization and Execution. Manuscript: Review and Critique.

ACKNOWLEDGMENTS

Authors want to acknowledge Mr. Simonluca Spadanuda for the creation of the random forest illustration (**Figure 1**).

REFERENCES

- Abraham, A., Milham, M. P., Di Martino, A., Craddock, R. C., Samaras, D., Thirion, B., et al. (2017). Deriving reproducible biomarkers from multi-site resting-state data: an autism-based example. *Neuroimage* 147, 736–745. doi: 10.1016/j.neuroimage.2016.10.045
- Acosta-Cabrero, J., and Nestor, P. J. (2014). Diffusion tensor imaging in Alzheimer's disease: insights into the limbic-diencephalic network and methodological considerations. *Front. Aging. Neurosci.* 6:266. doi: 10.3389/fnagi.2014.00266
- Ardekani, B. A., Bermudez, E., Mubeen, A. M., Bachman, A. H., and Alzheimer's Disease Neuroimaging, I. (2017). Prediction of incipient Alzheimer's disease dementia in patients with mild cognitive impairment. *J. Alzheimers Dis.* 55, 269–281. doi: 10.3233/JAD-160594
- Berchthold, N. C., and Cotman, C. W. (1998). Evolution in the conceptualization of dementia and Alzheimer's disease: Greco-Roman period to the 1960s. *Neurobiol. Aging* 19, 173–189. doi: 10.1016/S0197-4580(98)00052-9
- Borza, T., Engedal, K., Bergh, S., Benth, J., and Selbæk, G. (2015). The course of depression in late life as measured by the montgomery and asberg depression rating scale in an observational study of hospitalized patients. *BMC Psychiatry* 15:191. doi: 10.1186/s12888-015-0577-8
- Breiman, L. (1996). Bagging predictors. *Mach. Learn.* 24, 123–140. doi: 10.1007/BF00058655
- Breiman, L. (2001). Random Forests. *Mach. Learn.* 45, 5–32. doi: 10.1023/A:1010933404324
- Breiman, L., Friedman, J., Stone, C. J., and Olshen, R. A. (1984). *Classification and Regression Trees*. Boca Raton, FL: CRC press.
- Cabral, C., Silveira, M., and Alzheimer's Disease Neuroimaging, I. (2013). Classification of Alzheimer's disease from FDG-PET images using favourite class ensembles. *Conf. Proc. IEEE Eng. Med. Biol. Soc.* 2013, 2477–2480. doi: 10.1109/EMBC.2013.6610042
- Calle, M. L., Urrea, V., Boulesteix, A. L., and Malats, N. (2011). AUC-RF: a new strategy for genomic profiling with random forest. *Hum. Hered.* 72, 121–132. doi: 10.1159/000330778
- Caruana, R., and Niculescu-Mizil, A. (2006). "An empirical comparison of supervised learning algorithms," in *23rd International Conference on Machine Learning*, (Pittsburgh, PA: ACM Press), 161–168.
- Ceriani, L., and Verme, P. (2012). The origins of the gini index: extracts from variabilità e mutabilità (1912) by corrado gini. *J. Econ. Inequal.* 10, 421–443. doi: 10.1007/s10888-011-9188-x
- Chen, X., Wang, M., and Zhang, H. (2011). The use of classification trees for bioinformatics. *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.* 1, 55–63. doi: 10.1002/widm.14
- Collie, A., and Maruff, P. (2000). The neuropsychology of preclinical Alzheimer's disease and mild cognitive impairment. *Neurosci. Biobehav. Rev.* 24, 365–374. doi: 10.1016/S0149-7634(00)00012-9
- Dansereau, C., Benhajali, Y., Risterucci, C., Pich, E. M., Orban, P., Arnold, D., et al. (2017). Statistical power and prediction accuracy in multisite resting-state fMRI connectivity. *Neuroimage* 149, 220–232. doi: 10.1016/j.neuroimage.2017.01.072
- Falahati, F., Westman, E., and Simmons, A. (2014). Multivariate data analysis and machine learning in Alzheimer's disease with a focus on structural magnetic resonance imaging. *J. Alzheimers Dis.* 41, 685–708. doi: 10.3233/JAD-131928
- Fripp, J., Bourgeat, P., Acosta, O., Raniga, P., Modat, M., Pike, K. E., et al. (2008). Appearance modeling of 11C PiB PET images: characterizing amyloid deposition in Alzheimer's disease, mild cognitive impairment and healthy aging. *Neuroimage* 43, 430–439. doi: 10.1016/j.neuroimage.2008.07.053
- Frisoni, G. B., Fox, N. C., Jack, C. R. Jr., Scheltens, P., and Thompson, P. M. (2010). The clinical use of structural MRI in Alzheimer disease. *Nat. Rev. Neurol.* 6, 67–77. doi: 10.1038/nrneurol.2009.215
- Gray, K. R., Aljabar, P., Heckemann, R. A., Hammers, A., Rueckert, D., and Alzheimer's Disease Neuroimaging, I. (2013). Random forest-based similarity measures for multi-modal classification of Alzheimer's disease. *Neuroimage* 65, 167–175. doi: 10.1016/j.neuroimage.2012.09.065
- Greicius, M. D., Srivastava, G., Reiss, A. L., and Menon, V. (2004). Default-mode network activity distinguishes Alzheimer's disease from healthy aging: evidence from functional MRI. *Proc. Natl. Acad. Sci. U.S.A.* 101, 4637–4642. doi: 10.1073/pnas.0308627101
- Lebedev, A. V., Westman, E., Van Westen, G. J., Kramberger, M. G., Lundervold, A., Aarsland, D., et al. (2014). Random forest ensembles for detection and prediction of Alzheimer's disease with a good between-cohort robustness. *Neuroimage Clin.* 6, 115–125. doi: 10.1016/j.nicl.2014.08.023
- Lebedeva, A. K., Westman, E., Borza, T., Beyer, M. K., Engedal, K., Aarsland, D., et al. (2017). MRI-based classification models in prediction of mild cognitive impairment and dementia in late-life depression. *Front. Aging. Neurosci.* 9:13. doi: 10.3389/fnagi.2017.00013
- Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gotzsche, P. C., Ioannidis, J. P., et al. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. *BMJ* 339:b2700. doi: 10.1136/bmj.b2700
- Maggipinto, T., Bellotti, R., Amoroso, N., Diacono, D., Donvito, G., Lella, E., et al. (2017). DTI measurements for Alzheimer's classification. *Phys. Med. Biol.* 62, 2361–2375. doi: 10.1088/1361-6560/aa5d8e
- Matsuda, H. (2007). Role of neuroimaging in Alzheimer's disease, with emphasis on brain perfusion SPECT. *J. Nucl. Med.* 48, 1289–1300. doi: 10.2967/jnumed.106.037218
- Menze, B. H., Kelm, B. M., Masuch, R., Himmelreich, U., Bachert, P., Petrich, W., et al. (2009). A comparison of random forest and its Gini importance with standard chemometric methods for the feature selection and classification of spectral data. *BMC Bioinformatics* 10:213. doi: 10.1186/1471-2105-10-213
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., and Group, P. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ* 339:b2535. doi: 10.1136/bmj.b2535
- Moradi, E., Pepe, A., Gaser, C., Huttunen, H., Tohka, J., and Alzheimer's Disease Neuroimaging, I. (2015). Machine learning framework for early MRI-based Alzheimer's conversion prediction in MCI subjects. *Neuroimage* 104, 398–412. doi: 10.1016/j.neuroimage.2014.10.002
- Oppedal, K., Eftestøl, T., Engan, K., Beyer, M. K., and Aarsland, D. (2015). Classifying dementia using local binary patterns from different regions in magnetic resonance images. *Int. J. Biomed. Imaging* 2015:572567. doi: 10.1155/2015/572567
- Palmqvist, S., Hertz, J., Minthon, L., Wattmo, C., Zetterberg, H., Blennow, K., et al. (2012). Comparison of brief cognitive tests and CSF biomarkers in predicting Alzheimer's disease in mild cognitive impairment: six-year follow-up study. *PLoS ONE* 7:e38639. doi: 10.1371/journal.pone.0038639
- Rathore, S., Habes, M., Ifikhar, M. A., Shacklett, A., and Davatzikos, C. (2017). A review on neuroimaging-based classification studies and associated feature extraction methods for Alzheimer's disease and its prodromal stages. *Neuroimage* 155, 530–548. doi: 10.1016/j.neuroimage.2017.03.057
- Sarica, A., Cerasa, A., Valentino, P., Yeatman, J., Trotta, M., Barone, S., et al. (2017). The corticospinal tract profile in amyotrophic lateral sclerosis. *Hum. Brain Mapp.* 38, 727–739. doi: 10.1002/hbm.23412
- Shen, D., Wu, G., and Suk, H. I. (2017). Deep learning in medical image analysis. *Annu. Rev. Biomed. Eng.* 19, 221–248. doi: 10.1146/annurev-bioeng-071516-044442
- Sivapriya, T. R., Kamal, A. R., and Thangaiyah, P. R. (2015). Ensemble merit merge feature selection for enhanced multinomial classification in Alzheimer's dementia. *Comput. Math. Methods Med.* 2015:676129. doi: 10.1155/2015/676129
- Son, S. J., Kim, J., and Park, H. (2017). Structural and functional connective fingerprints in mild cognitive impairment and Alzheimer's disease patients. *PLoS ONE* 12:e0173426. doi: 10.1371/journal.pone.0173426
- Strobl, C., Boulesteix, A. L., and Augustin, T. (2007). Unbiased split selection for classification trees based on the gini index. *Comput. Stat. Data Anal.* 52, 483–501. doi: 10.1016/j.csda.2006.12.030
- Tripoliti, E. E., Fotiadis, D. I., and Argyropoulou, M. (2007). A supervised method to assist the diagnosis of Alzheimer's disease based on functional magnetic resonance imaging. *Conf. Proc. IEEE Eng. Med. Biol. Soc.* 2007, 3426–3429. doi: 10.1109/IEMBS.2007.4353067
- Trzepacz, P. T., Yu, P., Sun, J., Schuh, K., Case, M., Witte, M. M., et al. (2014). Comparison of neuroimaging modalities for the prediction of conversion from mild cognitive impairment to Alzheimer's dementia. *Neurobiol. Aging* 35, 143–151. doi: 10.1016/j.neurobiolaging.2013.06.018

- Vieira, S., Pinaya, W. H., and Mechelli, A. (2017). Using deep learning to investigate the neuroimaging correlates of psychiatric and neurological disorders: methods and applications. *Neurosci. Biobehav. Rev.* 74, 58–75. doi: 10.1016/j.neubiorev.2017.01.002
- Wang, P., Chen, K., Yao, L., Hu, B., Wu, X., Zhang, J., et al. (2016). Multimodal classification of mild cognitive impairment based on partial least squares. *J. Alzheimers Dis.* 54, 359–371. doi: 10.3233/JAD-160102
- Zhang, D., Shen, D., and Alzheimer's Disease Neuroimaging, I. (2012). Predicting future clinical changes of MCI patients using longitudinal and multimodal biomarkers. *PLoS ONE* 7:e33182. doi: 10.1371/journal.pone.0033182

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.

Copyright © 2017 Sarica, Cerasa and Quattrone. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Early Changes in Alpha Band Power and DMN BOLD Activity in Alzheimer's Disease: A Simultaneous Resting State EEG-fMRI Study

Katharina Brueggen^{1*†}, Carmen Fiala^{2†}, Christoph Berger³, Sina Ochmann², Claudio Babiloni^{4,5} and Stefan J. Teipel^{1,2}

¹ German Center for Neurodegenerative Diseases, Rostock, Germany, ² Department of Psychosomatic Medicine, University Medicine Rostock, Rostock, Germany, ³ Department of Psychiatry, Neurology, Psychosomatics, and Psychotherapy in Childhood and Adolescence, University of Rostock, Rostock, Germany, ⁴ Department of Physiology and Pharmacology "Vittorio Erspamer", University of Rome "La Sapienza", Rome, Italy, ⁵ Department of Neuroscience, IRCCS San Raffaele Pisana, Rome, Italy

OPEN ACCESS

Edited by:

Pedro Rosa-Neto,
McGill University, Canada

Reviewed by:

Paul Gerson Unschuld,
University of Zurich, Switzerland
Panteleimon Giannakopoulos,
Université de Genève, Switzerland

*Correspondence:

Katharina Brueggen
katharina.brueggen@dzne.de

[†] These authors have contributed
equally to this work.

Received: 05 May 2017

Accepted: 19 September 2017

Published: 06 October 2017

Citation:

Brueggen K, Fiala C, Berger C, Ochmann S, Babiloni C and Teipel SJ (2017) Early Changes in Alpha Band Power and DMN BOLD Activity in Alzheimer's Disease: A Simultaneous Resting State EEG-fMRI Study. *Front. Aging Neurosci.* 9:319. doi: 10.3389/fnagi.2017.00319

Simultaneous resting state functional magnetic resonance imaging (rsfMRI)–resting state electroencephalography (rsEEG) studies in healthy adults showed robust positive associations of signal power in the alpha band with BOLD signal in the thalamus, and more heterogeneous associations in cortical default mode network (DMN) regions. Negative associations were found in occipital regions. In Alzheimer's disease (AD), rsfMRI studies revealed a disruption of the DMN, while rsEEG studies consistently reported a reduced power within the alpha band. The present study is the first to employ simultaneous rsfMRI–rsEEG in an AD sample, investigating the association of alpha band power and BOLD signal, compared to healthy controls (HC). We hypothesized to find reduced positive associations in DMN regions and reduced negative associations in occipital regions in the AD group. Simultaneous resting state fMRI–EEG was recorded in 14 patients with mild AD and 14 HC, matched for age and gender. Power within the EEG alpha band (8–12 Hz, 8–10 Hz, and 10–12 Hz) was computed from occipital electrodes and served as regressor in voxel-wise linear regression analyses, to assess the association with the BOLD signal. Compared to HC, the AD group showed significantly decreased positive associations between BOLD signal and occipital alpha band power in clusters in the superior, middle and inferior frontal cortex, inferior temporal lobe and thalamus ($p < 0.01$, uncorr., cluster size ≥ 50 voxels). This group effect was more pronounced in the upper alpha sub-band, compared to the lower alpha sub-band. Notably, we observed a high inter-individual heterogeneity. Negative associations were only reduced in the lower alpha range in the hippocampus, putamen and cerebellum. The present study gives first insights into the relationship of resting-state EEG and fMRI characteristics in an AD sample. The results suggest that positive associations between alpha band power and BOLD signal in numerous regions, including DMN regions, are diminished in AD.

Keywords: Alzheimer's disease, alpha rhythm, electroencephalography, functional magnetic resonance imaging, default mode network

INTRODUCTION

In Alzheimer's disease (AD), resting state functional magnetic resonance imaging (rsfMRI), and resting state electroencephalography (rsEEG) have only been used separately to measure pathological changes. RsfMRI studies showed decreased activity (Greicius et al., 2004; Zhu et al., 2013; Li et al., 2015) and disrupted functional connectivity (Greicius et al., 2004; Zhang et al., 2009, 2010; Agosta et al., 2012; Koch et al., 2012; Weiler et al., 2014; Xia et al., 2014) in the default mode network (DMN) in AD. The DMN includes the anterior and posterior cingulate cortex, precuneus, medial prefrontal cortex, inferior parietal cortex, and hippocampal formation (Shulman et al., 1997; Raichle et al., 2001; Greicius et al., 2003; Buckner et al., 2008). Functionally, it has been associated with episodic memory (Mazoyer et al., 2001; Buckner et al., 2008; Weiler et al., 2014) and self-referential thinking (Raichle et al., 2001; Greicius et al., 2004; Buckner et al., 2008; Knyazev, 2013). Furthermore, rsEEG analyses showed reduced power within the alpha band (8–12 Hz) at early AD stages, as well as a slowing of the alpha rhythm and increased presence of lower frequency bands (Brenner et al., 1986; Dierks et al., 1993; Huang et al., 2000; Jeong, 2004). The alpha band is the dominant rhythm in healthy adults during a state of relaxed wakefulness, keeping the eyes closed (Berger, 1929; Zschocke and Hansen, 2012; Hinrichs, 2015). It originates from thalamo-cortical neurons projecting to the occipital cortex (Lorincz et al., 2009; Hughes et al., 2011; Zschocke and Hansen, 2012; Babiloni et al., 2015) – a projection pathway that may be disrupted in AD, as shown previously in studies using a computational model (Bhattacharya et al., 2011) and fMRI functional connectivity (Zhou et al., 2013). Functionally, alpha band power was shown to correlate positively with internal mental processes (Knyazev et al., 2011). Moreover, subdivisions of the alpha band may be related to different cognitive functions: the lower alpha band (8–10 Hz) may be associated with attention, while the upper alpha band (10–12 Hz) may be associated with memory processes (Klimesch, 1999). In addition, alpha band power has been suggested to play a role in an inhibitory gating mechanism of the visual system, suppressing unattended visual information (Berger, 1929; Palva and Palva, 2007; Tuladhar et al., 2007; Zumer et al., 2014). Power within the alpha band has been shown to correlate negatively with hemodynamic activity in the occipital cortex (Goldman et al., 2002; Moosmann et al., 2003; Gonçalves et al., 2006; Mantini et al., 2007; Scheeringa et al., 2012).

In order to assess the temporal association within subjects, the two modalities need to be measured simultaneously. The simultaneous rsfMRI-rsEEG measurement allows investigating the correlation of the BOLD signal fluctuation (as measured with rsfMRI) with the power fluctuation in specific frequency bands (as measured with rsEEG) over time. This method has previously been applied in young healthy subjects, correlating power fluctuations within the alpha band with BOLD signal fluctuations within each voxel. Most of these studies found that alpha band power fluctuation correlated positively with BOLD signal fluctuations in the thalamus (Goldman et al.,

2002; Moosmann et al., 2003; Gonçalves et al., 2006) and in cortical DMN regions (Mantini et al., 2007; Jann et al., 2009, 2010; Scheeringa et al., 2012). On the other hand, some studies reported only weak or no positive associations (Laufs et al., 2003a; Gonçalves et al., 2006; Mo et al., 2013). Negative associations were found between alpha band power fluctuation and BOLD signal fluctuation in occipital, parietal, and frontal cortical regions in young HC subjects (Goldman et al., 2002; Laufs et al., 2003a, 2006; Moosmann et al., 2003; Gonçalves et al., 2006; Mantini et al., 2007; Scheeringa et al., 2012).

The present study is the first to employ simultaneous fMRI-EEG measurement in AD patients. Its aim was to explore its feasibility and to investigate the relationship of alpha band power fluctuation and BOLD signal fluctuation in AD patients compared to HC subjects. As previous research showed alpha band power to correlate significantly with gray matter volume in AD (Babiloni et al., 2009, 2013, 2015), we additionally controlled for volume of the hippocampus, which is affected early in the disease (Devanand et al., 2007; den Heijer et al., 2010; Frisoni et al., 2010; Jack et al., 2011). We hypothesized to find positive associations between occipital alpha band power fluctuation and BOLD signal fluctuation in regions of the DMN in both groups (AD and HC), with a reduced association in the AD group. Secondly, we hypothesized to find positive associations of alpha band power fluctuation and BOLD signal fluctuation in the thalamus in both groups, but a weaker association in AD. Finally, we expected to find negative associations with BOLD signal fluctuation in the occipital cortex, with reduced associations in the AD group (Moretti, 2004).

MATERIALS AND METHODS

Participants

The groups consisted of $n = 14$ individuals each, matched for age and gender (see **Table 1** for demographic and clinical characteristics). Initially, $n = 18$ patients with mild AD and $n = 17$ elderly healthy control (HC) subjects participated in the study, of which one patient aborted the scan session, and three patients were excluded due to radiological abnormalities. Three female participants in the HC group were randomized out, in order to match the groups for gender. Patients were recruited via the memory clinic at the University Medicine Rostock (UMR); HC subjects were recruited via the database of the UMR, containing healthy subjects who were originally recruited via advertisement. HC were required to score within one standard deviation on all subscales of the Consortium to Establish a Registry for Alzheimer's Disease (CERAD) battery (Morris et al., 1989). Patients were clinically diagnosed with probable AD dementia according to the NINCDS-ADRDA and NIA-AA criteria (McKhann, 1984; McKhann et al., 2011). All subjects underwent general medical, neurological and psychiatric assessment. Neuropsychological assessment was conducted using the CERAD battery. Laboratory analyses and APOE genotype sequencing were carried out. Subjects exhibited no neurological or radiological abnormalities (e.g., normal pressure hydrocephalus or extensive microinfarcts), and no

psychiatric diseases. AD patients showed no signs of dementia not due to AD (e.g., vascular dementia). The study was approved by the local ethics committee of the University Rostock. All participants gave written informed consent, and all procedures were carried out in accordance with the Helsinki declaration.

Data Acquisition

Electroencephalography and fMRI data were recorded simultaneously during 7.5 min of resting state (eyes-closed). For the EEG recording, MRI-compatible measurement devices (Brain Products, Gilching, Germany) and the software Brain Vision Recorder¹ were used. EEG was recorded at 32 electrodes that were positioned according to the international 10-20-system (Jasper, 1958). The reference electrode was located between Fz and Cz, the ground electrode at AFz. Impedances of the electrodes of interest (O1, O2, and Oz) were kept below 8 k Ω , except for one AD patient (18 k Ω). An additional ECG channel was attached to detect cardio-ballistic artifacts. EEG data were sampled at 5 kHz. The EEG amplifier sampling interval was phase-synchronized to the fMRI main frequency via the Syncbox (Brain Products, Gilching, Germany) in order to preclude EEG-fMRI-sampling-jitter artifacts. The EEG hardware (i.e., amplifier and powerpack) was placed at the head end of the scanner tube and weighted with sand bags to prevent hardware motion.

Functional magnetic resonance imaging images were acquired using a 3-Tesla Siemens Magnetom scanner with a T2-weighted echo-planar imaging sequence (TR: 2.6 s, TE: 30 ms, FOV: 224 mm, thickness: 3.5 mm, number of slices: 180). The anatomical images were recorded using a T1-weighted MPRAGE sequence (TR: 2.5 s, TE: 4.37 ms, FOV 256 mm, thickness: 1 mm, number of slices: 192). Foam wedges were used to stabilize the head. Subjects were instructed to stay awake, keeping their eyes closed. The EEG signal was visually controlled for signs of sleep (offline).

Data Preprocessing

EEG Data

Data were preprocessed using Brain Vision Analyzer software (Version 2.0, Brain Products, Gilching, Germany). First, data were downsampled to 250 Hz. Imaging and ECG pulse artifacts

were eliminated using the average artifact subtraction method described by Allen et al. (1998, 2000), which is included in the Brain Vision Analyzer software. Briefly, the imaging artifacts were automatically marked based on recurring patterns, the thus-defined intervals were averaged and their means subtracted from each interval. ECG pulse artifacts were removed by constructing an average ECG artifact template and subtracting it from the EEG data. Data were high-pass (0.5 Hz) and low-pass (70 Hz) filtered. Additionally, a notch filter was applied at 50 Hz. Using Independence Component Analysis (ICA), artifacts caused by eye movement, temporal electrode noise and residual pulse artifacts were removed. In case the electrode noise could not be eliminated by removing two independent components, the disturbed channel was removed and interpolated by topographical triangulation (occipital channels were not affected by this). After ICA, the data were again visually inspected for residual artifacts. No sleep patterns (i.e., K-complexes or sleep spindles) were present. EEG data from the AD group showed more artifacts such as eye movement and muscle activation, especially during the second half of the scan time, possibly constituting a sign of growing unrest. Two AD subjects showed a shift in frequency from alpha to theta over time. These artifacts were removed. The EEG signal was re-referenced to a common reference, obtained by averaging across all channels.

The electrodes O1, O2, and Oz were chosen as electrodes of interest, since alpha activity is best expressed at occipital electrodes (Moosmann et al., 2003; Laufs et al., 2003a; Mo et al., 2013). The arithmetic mean of electrophysiological activity from O1, O2, and Oz was calculated. Using complex demodulation, the EEG time courses of power within the total (8–12 Hz), lower (8–10 Hz), and upper (10–12 Hz) alpha band were extracted for each individual and exported to MATLAB (Mathworks, Inc., Sherborn, MA, United States) for the creation of statistical model regressors.

MRI Data

Functional magnetic resonance imaging data preprocessing was performed using SPM8² implemented in Matlab 7 (Mathworks, Natick) and the VBM8 toolbox (Version 414³). The first six volumes were removed to eliminate saturation effects. Slices were referenced to the temporally middle slice. After realignment of the functional images, the anatomical images were coregistered to the realigned mean functional image. The structural T1-weighted MPRAGE images were segmented into gray matter, white matter and cerebrospinal fluid compartments and warped to standard MNI space, using the default MNI standard template and the Diffeomorphic Anatomical Registration Through Exponentiated Lie Algebra (DARTEL) method (Ashburner, 2007) implemented in VBM8. The resulting deformation fields were used to warp the functional images to standard space. Spatial smoothing of the normalized functional images was performed with a Gaussian Kernel of 8 mm full-width half-maximum (FWHM). In order to reduce slow drift artifacts, a high-pass filter with a cut-off period of 128 s was applied to the voxel time

¹www.brainproducts.com

TABLE 1 | Demographic and clinical characteristics of the study subjects; mean \pm SD (range).

| | AD (n = 14) | HC (n = 14) | p* |
|---|------------------------|------------------------|--------|
| Age | 75.3 \pm 5.7 (64–82) | 73.4 \pm 3.1 (68–79) | 0.276 |
| Gender (male/female) | 10/4 | 10/4 | n. a. |
| Education (years) | 14.4 \pm 2.7 (8–17) | 13.6 \pm 2.8 (11–20) | 0.417 |
| MMSE score | 24.6 \pm 3.1 (17–28) | 28.7 \pm 0.8 (27–30) | <0.001 |
| APOE status (E2/E3; E2/E4; E3/E3; E3/E4; E4/E4) | 0; 2; 4; 6; 1 | 2; 1; 7; 2; 1 | n. a. |

*Independent samples t-test, 2-sided.

²<http://www.fil.ion.ucl.ac.uk/spm/>

³<http://dbm.neuro.uni-jena.de/vbm8/>

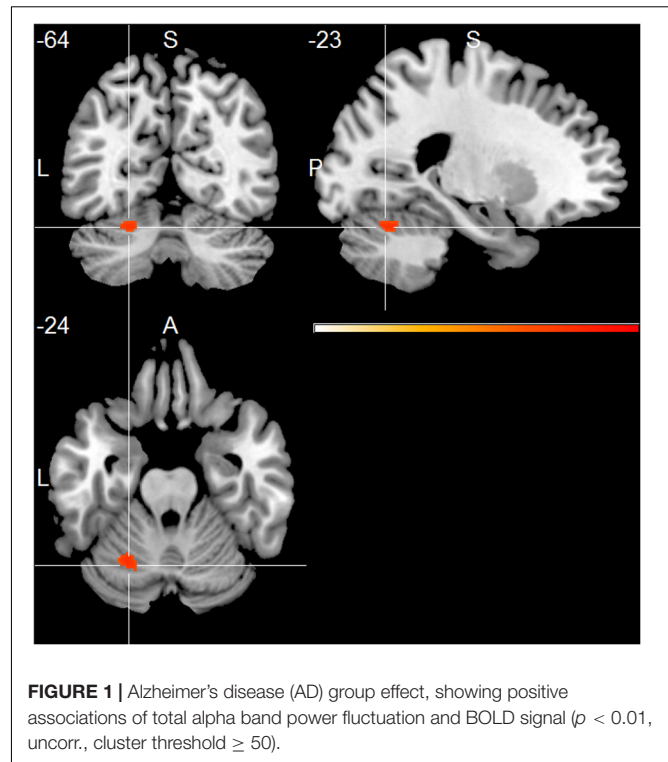
courses. From the segmented gray matter images, gray matter volume of the left and right hippocampus was calculated for each subject, using binarized inclusive masks that had been created for the IXI template in MNI space according to the international harmonization protocol for hippocampus segmentation (Grothe et al., 2012; Boccardi et al., 2015). The volume of the left and right hippocampus was pooled and normalized by dividing it by the total intracranial volume.

A regressor containing one-second intervals of artifact-free, averaged spectral power of the pooled occipital electrodes and an additional on/off regressor of no interest (containing timing information of artifacts longer than 1 s) were created. Separate regressors were built for power within the total alpha (8–12 Hz), lower alpha (8–10 Hz) and upper alpha band (10–12 Hz). The regressors of interest were convolved with an *a priori* defined hemodynamic response function (HRF) (Cohen, 1997) within the SPM first-level (single subject) processing pipeline (for a diagram see Supplementary Figure 1).

Statistical Analysis

For comparing relative alpha band power at the pooled occipital channels (O1, O2, and Oz) between groups, Fast Fourier transformation (FFT) across 1-s-segments was used. Two-sided independent samples *t*-tests were used to compare relative alpha power and normalized hippocampal gray matter volume between groups. Separate general linear models were specified for total alpha, lower alpha and upper alpha, respectively, using SPM8 (Friston et al., 2007). The models included a regressor variable containing the power information for the respective HRF-convolved alpha band, a mean term regressor, a covariate regressor containing the artifact information, and the covariates age, gender, and years of education. For the first-level analysis, positive and negative *t*-contrasts were specified for each subject, testing for the effects of the alpha band power regressor, controlled for the artifact regressor. This resulted in individual statistical parametric maps of positive and of negative associations of the total, lower or upper alpha power fluctuation over time, respectively, with the BOLD fluctuation in each voxel of the brain. The resulting maps of EEG regressor weights were used for group comparisons in one- and two-sample *t*-tests. The one-sample *t*-tests were performed for the AD and HC group separately, testing for positive and negative associations of each alpha regressor weight across all subjects in the respective group. For the two-sample *t*-test, a contrast of HC > AD was defined for positive and negative associations, respectively. The second-level analyses were additionally controlled for the covariate regressor normalized hippocampal gray matter volume.

All statistical results were restricted to voxels within gray matter, by thresholding the default IXI template in VBM8 at $p < 0.3$ and using it as inclusive mask. Statistical significance levels were set at $p < 0.01$ (uncorrected for multiple comparisons). Only clusters with a voxel count ≥ 50 were considered. Resulting clusters were visually compared to a functional connectivity based DMN atlas (Shirer et al., 2012).



RESULTS

Alpha Power Fluctuations

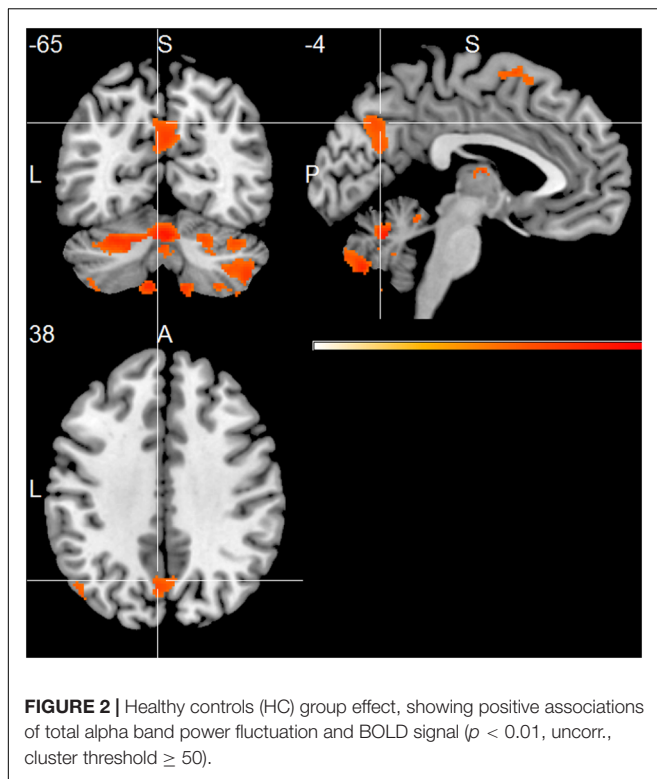
The mean relative alpha band power was not significantly different between the groups (AD: $35.0 \pm 17.7\%$; HC: $32.0 \pm 21.6\%$, Supplementary Table 1). However, at visual inspection, a morphological difference in the form of dysmorphic alpha waves was observed in the AD group.

Association of Alpha Band Power and fMRI BOLD Dynamics

Positive Associations

At group level, the AD group showed positive associations of total alpha band power with BOLD fluctuation in the cerebellum (one sample *t*-test, $p < 0.01$, uncorr., **Figure 1** and Supplementary Table 2). Lower alpha band power correlated positively with clusters in the right inferior temporal lobe, right hippocampus, left putamen and cerebellum ($p < 0.01$, uncorr.) (Supplementary Table 2). In contrast, power within the upper alpha frequency showed no significant positive associations in any regions.

The HC group showed positive associations of total alpha band power with mainly frontal and temporal cortical regions, including superior, middle and inferior frontal cortex, temporal pole, parietal cortex, thalamus, putamen and cerebellum (one sample *t*-test, $p < 0.01$, uncorr., **Figure 2** and Supplementary Table 3). Within the lower alpha frequency, fewer associations were present, which were located mainly in frontal regions, left inferior temporal lobe, thalamus and cerebellum. Most associations were found within the upper frequency, located



mainly in the hippocampus, thalamus, occipital, temporal and frontal cortex, including anterior cingulate cortex and middle cingulate, putamen and caudate nucleus, as well as cerebellum (Supplementary Table 3).

Compared to the HC group, the AD group showed significantly decreased positive associations of total alpha band power with BOLD fluctuation in clusters in the frontal cortex (superior, middle, inferior, precentral gyrus, and anterior cingulate cortex), inferior temporal lobe and thalamus (two-samples t -test, $p < 0.01$, uncorr., **Figure 3** and Supplementary Table 4). Similar decreased associations were found for the upper alpha band power (superior frontal lobe, insula and parietal lobe) (**Figure 4** and Supplementary Table 4). Regarding the lower alpha band power, the AD group showed decreased positive associations with scattered clusters in the superior frontal lobe, compared to the HC group (Supplementary Table 4).

At the individual level, first-level analyses revealed positive associations of power within the total alpha band range with regions that belong to the DMN (Shirer et al., 2012) in $n = 6$ HC subjects and in $n = 3$ AD patients (**Table 2**). For an example, see Supplementary Figures 2, 3.

Normalized hippocampal gray matter volume was lower in the AD group, although not significantly (independent samples t -test; $T(26) = 1.735$, $p = 0.095$). Entering it as covariate regressor in the general linear models did not essentially change the results of the one- and two-sample t -tests (Supplementary Figures 4–6).

Negative Associations

At group level, the AD group showed negative associations of total band alpha power with clusters in the occipital, frontal

and temporal cortex (one-sample t -test, $p < 0.01$, uncorr., Supplementary Table 5). In the upper alpha band, associations were only significant in the occipital cortex. Lower alpha band power showed no significant associations (Supplementary Table 5).

The HC group showed significant negative associations of total alpha band power with clusters in the precentral gyrus and superior temporal cortex (one-sample t -test, $p < 0.01$, uncorr., Supplementary Table 6). No suprathreshold clusters were found in the upper alpha band. Lower alpha band power showed pronounced negative associations with the frontal cortex, mainly in the precentral and paracentral gyrus, and with the parietal cortex, temporal and middle cingulate cortex (Supplementary Table 6).

Compared to the HC group, the AD group did not exhibit significantly reduced negative associations of total or upper alpha band power with BOLD signal fluctuation in any voxel clusters. Regarding the lower alpha band, significantly decreased negative associations were found in the hippocampus, putamen and cerebellum (two-sample t -test, $p < 0.01$, uncorr., Supplementary Table 7).

At the individual level, first-level analyses revealed negative associations of alpha band power with BOLD fluctuations in both anterior and posterior regions in $n = 5$ AD patients and $n = 7$ HC subjects, associations in mainly frontal regions in $n = 3$ AD patients and $n = 2$ HC subjects, and associations in mainly posterior regions in $n = 1$ HC subject.

DISCUSSION

The study successfully applied simultaneous fMRI-EEG to an AD sample for the first time and showed a reduced positive association between alpha band power and BOLD fluctuations in the AD patients, compared to the HC subjects. In the HC group, positive associations between alpha band power and BOLD fluctuations were observed in numerous regions, including DMN regions. Although present in all alpha sub-bands, they were especially evident in the upper alpha frequency band. The reduction of these positive associations in the AD patients might be due to altered functional interaction between the brain regions (Greicius et al., 2004; Zhang et al., 2009, 2010; Agosta et al., 2012; Weiler et al., 2014; Xia et al., 2014). The functional associations were not altered by the correction for hippocampal volume, indicating that they were not driven by atrophy.

Based on previous simultaneous fMRI-EEG studies with healthy participants, we hypothesized to find a positive association of alpha band power and BOLD signal fluctuation in the thalamus in HC subjects (Goldman et al., 2002; Moosmann et al., 2003; Gonçalves et al., 2006). In the light of the disrupted integrity of the thalamo-cortical system, we expected this association to be reduced in the AD patients (Bhattacharya et al., 2011; Zhou et al., 2013). In line with the hypothesis, these associations were present in the HC group and were decreased in the AD group. Additionally, in both groups, we found more positive associations of the upper alpha band

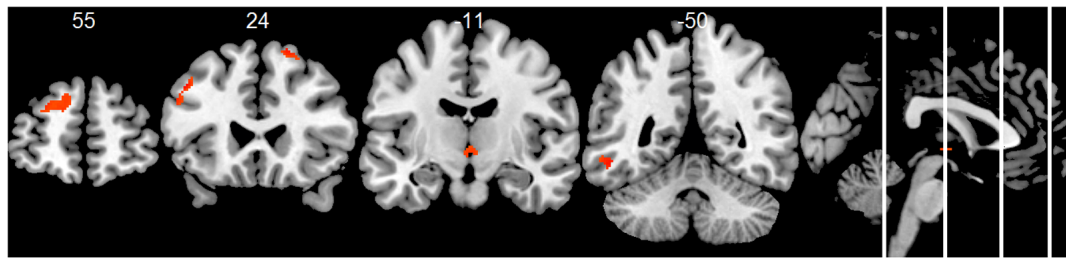


FIGURE 3 | Group comparison HC > AD of positive associations of total alpha band power fluctuation and BOLD signal ($p < 0.01$, uncorr., cluster threshold ≥ 50).

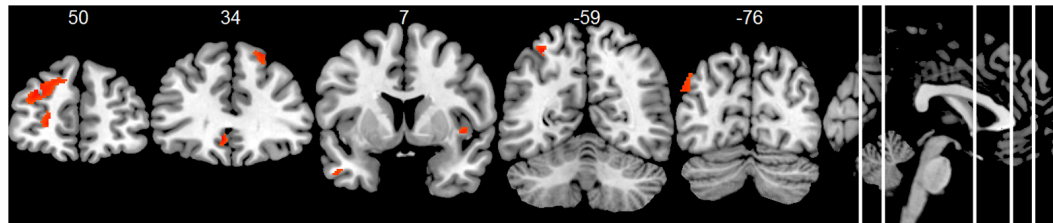


FIGURE 4 | Group comparison HC > AD of positive associations of upper alpha band power fluctuation and BOLD signal ($p < 0.01$, uncorr., cluster threshold ≥ 50).

power with the thalamus compared to the lower alpha band. This might indicate a frequency-specificity. Also, as thalamo-cortical activity underlies alpha generation and modulation (Bhattacharya et al., 2013), future functional connectivity studies might investigate whether decreased associations of alpha band power and thalamic BOLD fluctuations are related to the thalamo-cortical connectivity in AD (Zhou et al., 2013).

The third hypothesis included finding negative associations with BOLD signal fluctuation in the occipital cortex. Negative associations were found at group level in AD patients in the occipital cortex, as well as superior medial frontal cortex and temporal cortex. However, we did not find negative associations with the occipital cortex in HC subjects at group level. This is in contrast to a number of fMRI-EEG studies in young healthy subjects, showing negative associations of alpha band power with BOLD signal in the occipital cortex (Goldman et al., 2002;

Moosmann et al., 2003; Gonçalves et al., 2006; Mantini et al., 2007; Scheeringa et al., 2012). In the light of the overall accepted theory that alpha band represents a hallmark of the resting state of the brain (e.g., Gonçalves et al., 2006), we would have expected it to correlate negatively with BOLD signaling in the respective region. Instead, we found negative associations at HC group level in frontal, temporal and parietal regions. Although unexpected, this result is in line with a few other studies that reported an absence of negative associations with BOLD signal in the occipital cortex (Laufs et al., 2003a,b; Jann et al., 2009).

Interestingly, positive as well as negative associations with the cerebellum were present in almost all subjects. The cerebellum has received little attention in previous fMRI-EEG research (Scheeringa et al., 2012). FMRI studies showed impaired functional connectivity of the cerebellum in AD (Zheng et al., 2017), and a sensitivity of the cortico-cerebellar coupling to amyloid- β load in HC (Steininger et al., 2014). It would be interesting for future research to investigate the association of alpha band power and the integrity of cortical-cerebellar functional processes during rest.

A general limitation of fMRI resting state measurement is its high variability over time (Cole et al., 2010; Chen et al., 2015). The instruction to keep the eyes closed and to stay awake leaves room for spontaneous cognitive processes with varying attentional states. Possibly, the activation of the DMN might have been more robust if a task-based study design had been used, for example involving tasks of self-referential thinking or autobiographical memory (Andreasen et al., 1995; Mitchell, 2006; Gobbi et al., 2007; Spreng and Grady, 2010; Knyazev et al., 2011; Fomina et al., 2015). However, to be able to draw inferences on a potential clinical use, a resting state paradigm was needed. Another limitation is the relatively liberal statistical threshold. As this was the first study to employ simultaneous rsfMRI-rsEEG in

TABLE 2 | First-level analyses: number of subjects (n) showing positive associations of alpha band power and BOLD signal fluctuations, significant at $p < 0.01$ (uncorr.).

| | Default mode network* | Thalamus |
|------------------------|-----------------------|----------|
| AD | | |
| Total alpha (8–12 Hz) | 3 | 3 |
| Lower alpha (8–10 Hz) | 3 | 2 |
| Upper alpha (10–12 Hz) | 4 | 3 |
| HC | | |
| Total alpha (8–12 Hz) | 6 | 5 |
| Lower alpha (8–10 Hz) | 4 | 3 |
| Upper alpha (10–12 Hz) | 7 | 5 |

*Encompassing three or more of the following regions: precuneus, PCC, ACC, medial prefrontal cortex, and inferior parietal lobe.

AD patients, we aimed to assess the feasibility and to explore the associations in the whole brain.

We noted a high regional variability of both positive and negative associations between alpha band power fluctuation and BOLD signal between individual subjects, which has also been reported in previous studies (Goldman et al., 2002; Gonçalves et al., 2006; Laufs et al., 2006). Variability has been suggested to be partly caused by fluctuations in vigilance (Goldman et al., 2002; Laufs et al., 2006). Although our data were visually controlled for sleep, fluctuations in vigilance may have been present, particularly as an increase in artifacts in AD patients toward the end of the scan time was noted. The effect of vigilance on the association patterns of rsEEG and rsfMRI should be addressed in future research. Our results of high inter-individual heterogeneity, taken together with findings of high inter- and intra-individual variability observed in other resting state fMRI-EEG studies (Goldman et al., 2002; Laufs et al., 2003a, 2006; Moosmann et al., 2003; Gonçalves et al., 2006; Jann et al., 2009; Olbrich et al., 2009), also highlight the importance of future research with larger samples to be able to identify subgroups. Furthermore, our results support the necessity to differentiate the alpha band into sub-bands, as more HC subjects showed positive association patterns within the upper sub-band. This agrees with some other studies that investigated separate sub-bands (Laufs et al., 2006; Jann et al., 2009, 2010), linking sub-bands to different cognitive functions (e.g., Klimesch, 1999) and even indicating the possibility of predicting conversion from MCI to AD by calculating the ratio of power in alpha sub-bands (Moretti, 2015).

CONCLUSION

The present study showed diminished positive associations between alpha band power fluctuation and BOLD signal fluctuations in several brain regions in AD patients, compared

to HC subjects. These regions included (but were not limited to) DMN and thalamic regions. This study demonstrates the feasibility of measuring simultaneous rsEEG and rsfMRI signal fluctuations in a clinical AD population. Further research is needed to corroborate and expand its results.

AUTHOR CONTRIBUTIONS

KB recruited participants, performed neuropsychological testing, acquired EEG and MRI data, performed preprocessing and analyses, interpreted the data, drafted and revised the manuscript. CF recruited participants, conducted physical examinations, acquired EEG and MRI data, performed preprocessing and analyses, interpreted the data, and drafted the manuscript. CBE performed preprocessing and statistical analyses, interpreted the data, and revised the manuscript. SO contributed to the data interpretation and was involved in drafting the manuscript. CBA contributed to the study design, provided intellectual content for data interpretation, and revised the manuscript. ST was involved in all stages of the study, establishing the study design, recruiting participants, performing physical examinations, and revising the manuscript.

FUNDING

This study was supported by a grant of the Federal Ministry of Research (BMBF) to ST (AgeGain, 01GQ1425B).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fnagi.2017.00319/full#supplementary-material>

REFERENCES

- Agosta, F., Pievani, M., Geroldi, C., Copetti, M., Frisoni, G. B., and Filippi, M. (2012). Resting state fMRI in Alzheimer's disease: beyond the default mode network. *Neurobiol. Aging* 33, 1564–1578. doi: 10.1016/j.neurobiolaging.2011.06.007
- Allen, P. J., Josephs, O., and Turner, R. (2000). A method for removing imaging artifact from continuous EEG recorded during functional MRI. *Neuroimage* 12, 230–239. doi: 10.1006/nimg.2000.0599
- Allen, P. J., Polizzi, G., Krakow, K., Fish, D. R., and Lemieux, L. (1998). Identification of EEG events in the MR scanner: the problem of pulse artifact and a method for its subtraction. *Neuroimage* 8, 229–239. doi: 10.1006/nimg.1998.0361
- Andreasen, N. C., O'Leary, D. S., Cizadlo, T., Arndt, S., Rezai, K., Watkins, G. L., et al. (1995). Remembering the past: two facets of episodic memory explored with positron emission tomography. *Am. J. Psychiatry* 152, 1576–1585. doi: 10.1176/ajp.152.11.1576
- Ashburner, J. (2007). A fast diffeomorphic image registration algorithm. *Neuroimage* 38, 95–113. doi: 10.1016/j.neuroimage.2007.07.007
- Babiloni, C., Carducci, F., Lizio, R., Vecchio, F., Baglieri, A., Bernardini, S., et al. (2013). Resting state cortical electroencephalographic rhythms are related to gray matter volume in subjects with mild cognitive impairment and Alzheimer's disease. *Hum. Brain Mapp.* 34, 1427–1446. doi: 10.1002/hbm.22005
- Babiloni, C., Del Percio, C., Boccardi, M., Lizio, R., Lopez, S., Carducci, F., et al. (2015). Occipital sources of resting-state alpha rhythms are related to local gray matter density in subjects with amnesic mild cognitive impairment and Alzheimer's disease. *Neurobiol. Aging* 36, 556–570. doi: 10.1016/j.neurobiolaging.2014.09.011
- Babiloni, C., Frisoni, G. B., Pievani, M., Vecchio, F., Lizio, R., Buttligione, M., et al. (2009). Hippocampal volume and cortical sources of EEG alpha rhythms in mild cognitive impairment and Alzheimer disease. *Neuroimage* 44, 123–135. doi: 10.1016/j.neuroimage.2008.08.005
- Berger, H. (1929). Über das Elektrenkephalogramm des Menschen. *Arch. Psychiatr. Nervenkr.* 87, 527–570. doi: 10.1007/BF01797193
- Bhattacharya, B. S., Cakir, Y., Serap-Sengor, N., Maguire, L., and Coyle, D. (2013). Model-based bifurcation and power spectral analyses of thalamocortical alpha rhythm slowing in Alzheimer's disease. *Neurocomputing* 115, 11–22. doi: 10.1016/j.neucom.2012.10.023
- Bhattacharya, B. S., Coyle, D., and Maguire, L. P. (2011). A thalamo-cortico-thalamic neural mass model to study alpha rhythms in Alzheimer's disease. *Neural Netw.* 24, 631–645. doi: 10.1016/j.neunet.2011.02.009
- Boccardi, M., Bocchetta, M., Apostolova, L. G., Barnes, J., Bartzokis, G., Corbetta, G., et al. (2015). Delphi definition of the EADC-ADNI Harmonized Protocol for hippocampal segmentation on magnetic resonance. *Alzheimers Dement.* 11, 126–138. doi: 10.1016/j.jalz.2014.02.009

- Brenner, R. P., Ulrich, R. F., Spiker, D. G., Scabassi, R. J., Reynolds, C. F., Marin, R. S., et al. (1986). Computerized EEG spectral analysis in elderly normal, demented and depressed subjects. *Electroencephalogr. Clin. Neurophysiol.* 64, 483–492. doi: 10.1016/0013-4694(86)90184-7
- Buckner, R. L., Andrews-Hanna, J. R., and Schacter, D. L. (2008). The brain's default network: anatomy, function, and relevance to disease. *Ann. N. Y. Acad. Sci.* 1124, 1–38. doi: 10.1196/annals.1440.011
- Chen, B., Xu, T., Zhou, C., Wang, L., Yang, N., Wang, Z., et al. (2015). Individual variability and test-retest reliability revealed by ten repeated resting-state brain scans over one month. *PLOS ONE* 10:e0144963. doi: 10.1371/journal.pone.0144963
- Cohen, M. S. (1997). Parametric analysis of fMRI data using linear systems methods. *Neuroimage* 6, 93–103. doi: 10.1006/nimg.1997.0278
- Cole, D. M., Smith, S. M., and Beckmann, C. F. (2010). Advances and pitfalls in the analysis and interpretation of resting-state FMRI data. *Front. Syst. Neurosci.* 4:8. doi: 10.3389/fnsys.2010.00008
- den Heijer, T., van der Lijn, F., Koudstaal, P. J., Hofman, A., van der Lugt, A., Krestin, G. P., et al. (2010). A 10-year follow-up of hippocampal volume on magnetic resonance imaging in early dementia and cognitive decline. *Brain* 133, 1163–1172. doi: 10.1093/brain/awq048
- Devanand, D. P., Pradhaban, G., Liu, X., Khandji, A., Santi, S. D., Segal, S., et al. (2007). Hippocampal and entorhinal atrophy in mild cognitive impairment: prediction of Alzheimer disease. *Neurology* 68, 828–836. doi: 10.1212/01.wnl.0000256697.20968.d7
- Dierks, T., Ihl, R., Frölich, L., and Maurer, K. (1993). Dementia of the Alzheimer type: effects on the spontaneous EEG described by dipole sources. *Psychiatry Res.* 50, 151–162. doi: 10.1016/0925-4927(93)90027-F
- Fomina, T., Hohmann, M., Scholkopf, B., and Grosse-Wentrup, M. (2015). Identification of the default mode network with electroencephalography. *Conf. Proc. IEEE Eng. Med. Biol. Soc.* 2015, 7566–7569. doi: 10.1109/EMBC.2015.7320143
- Frisoni, G. B., Fox, N. C., Jack, C. R., Scheltens, P., and Thompson, P. M. (2010). The clinical use of structural MRI in Alzheimer disease. *Nat. Rev. Neurol.* 6, 67–77. doi: 10.1038/nrneurol.2009.215
- Friston, K. J., Ashburner, J., Kiebel, S. J., Nichols, T., and Penny, W. D. eds (2007). *Statistical Parametric Mapping. The Analysis of Functional Brain Images*. London: Academic Press.
- Gobbini, M. I., Koralek, A. C., Bryan, R. E., Montgomery, K. J., and Haxby, J. V. (2007). Two takes on the social brain: a comparison of theory of mind tasks. *J. Cogn. Neurosci.* 19, 1803–1814. doi: 10.1162/jocn.2007.19.11.1803
- Goldman, R. I., Stern, J. M., Engel, J., and Cohen, M. S. (2002). Simultaneous EEG and fMRI of the alpha rhythm. *Neuroreport* 13, 2487–2492. doi: 10.1097/00001756-200212200-00022
- Gonçalves, S. I., Munck, J. C. D., Pouwels, P. J. W., Schoonhoven, R., Kuijer, J. P. A., Maurits, N. M., et al. (2006). Correlating the alpha rhythm to BOLD using simultaneous EEG/fMRI: inter-subject variability. *Neuroimage* 30, 203–213. doi: 10.1016/j.neuroimage.2005.09.062
- Greicius, M. D., Krasnow, B., Reiss, A. L., and Menon, V. (2003). Functional connectivity in the resting brain: a network analysis of the default mode hypothesis. *Proc. Natl. Acad. Sci. U.S.A.* 100, 253–258. doi: 10.1073/pnas.0135058100
- Greicius, M. D., Srivastava, G., Reiss, A. L., and Menon, V. (2004). Default-mode network activity distinguishes Alzheimer's disease from healthy aging: evidence from functional MRI. *Proc. Natl. Acad. Sci. U.S.A.* 101, 4637–4642. doi: 10.1073/pnas.0308627101
- Grothe, M., Heinsen, H., and Teipel, S. J. (2012). Atrophy of the cholinergic Basal forebrain over the adult age range and in early stages of Alzheimer's disease. *Biol. Psychiatry* 71, 805–813. doi: 10.1016/j.biopsych.2011.06.019
- Hinrichs, H. (2015). Das EEG. *Das Neurophysiol. Labor* 37, 12–21. doi: 10.1016/j.neulab.2014.10.013
- Huang, C., Wahlund, L. O., Dierks, T., Julin, P., Winblad, B., and Jelic, V. (2000). Discrimination of Alzheimer's disease and mild cognitive impairment by equivalent EEG sources: a cross-sectional and longitudinal study. *Clin. Neurophysiol.* 111, 1961–1967. doi: 10.1016/S1388-2457(00)00454-5
- Hughes, S. W., Lorincz, M. L., Blethyn, K., Kekesi, K. A., Juhasz, G., Turmaine, M., et al. (2011). Thalamic gap junctions control local neuronal synchrony and influence macroscopic oscillation amplitude during EEG alpha rhythms. *Front. Psychol.* 2:193. doi: 10.3389/fpsyg.2011.00193
- Jack, C. R., Albert, M. S., Knopman, D. S., McKhann, G. M., Sperling, R. A., Carrillo, M. C., et al. (2011). Introduction to the recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement.* 7, 257–262. doi: 10.1016/j.jalz.2011.03.004
- Jann, K., Dierks, T., Boesch, C., Kottlow, M., Strik, W., and Koenig, T. (2009). BOLD correlates of EEG alpha phase-locking and the fMRI default mode network. *Neuroimage* 45, 903–916. doi: 10.1016/j.neuroimage.2009.01.001
- Jann, K., Kottlow, M., Dierks, T., Boesch, C., Koenig, T., and Brembs, B. (2010). Topographic electrophysiological signatures of fMRI resting state networks. *PLOS ONE* 5:e12945. doi: 10.1371/journal.pone.0012945
- Jasper, H. H. (1958). The ten-twenty electrode system of the international federation. The international federation of clinical neurophysiology. *Electroencephalogr. Clin. Neurophysiol.* 52, 371–375.
- Jeong, J. (2004). EEG dynamics in patients with Alzheimer's disease. *Clin. Neurophysiol.* 115, 1490–1505. doi: 10.1016/j.clinph.2004.01.001
- Klimesch, W. (1999). EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. *Brain Res. Brain Res. Rev.* 29, 169–195. doi: 10.1016/S0165-0173(98)00056-3
- Knyazev, G. G. (2013). EEG correlates of self-referential processing. *Front. Hum. Neurosci.* 7:264. doi: 10.3389/fnhum.2013.00264
- Knyazev, G. G., Slobodskoj-Plusnin, J. Y., Bocharov, A. V., and Pylkova, L. V. (2011). The default mode network and EEG alpha oscillations: an independent component analysis. *Brain Res.* 1402, 67–79. doi: 10.1016/j.brainres.2011.05.052
- Koch, W., Teipel, S., Mueller, S., Benninghoff, J., Wagner, M., Bokde, A. L. W., et al. (2012). Diagnostic power of default mode network resting state fMRI in the detection of Alzheimer's disease. *Neurobiol. Aging* 33, 466–478. doi: 10.1016/j.neurobiolaging.2010.04.013
- Laufs, H., Holt, J. L., Elfont, R., Krams, M., Paul, J. S., Krakow, K., et al. (2006). Where the BOLD signal goes when alpha EEG leaves. *Neuroimage* 31, 1408–1418. doi: 10.1016/j.neuroimage.2006.02.002
- Laufs, H., Kleinschmidt, A., Beyerle, A., Eger, E., Salek-Haddadi, A., Preibisch, C., et al. (2003a). EEG-correlated fMRI of human alpha activity. *Neuroimage* 19, 1463–1476.
- Laufs, H., Krakow, K., Sterzer, P., Eger, E., Beyerle, A., Salek-Haddadi, A., et al. (2003b). Electroencephalographic signatures of attentional and cognitive default modes in spontaneous brain activity fluctuations at rest. *Proc. Natl. Acad. Sci. U.S.A.* 100, 11053–11058.
- Li, H.-J., Hou, X.-H., Liu, H.-H., Yue, C.-L., He, Y., and Zuo, X.-N. (2015). Toward systems neuroscience in mild cognitive impairment and Alzheimer's disease: a meta-analysis of 75 fMRI studies. *Hum. Brain Mapp.* 36, 1217–1232. doi: 10.1002/hbm.22689
- Lorincz, M. L., Kekesi, K. A., Juhasz, G., Crunelli, V., and Hughes, S. W. (2009). Temporal framing of thalamic relay-mode firing by phasic inhibition during the alpha rhythm. *Neuron* 63, 683–696. doi: 10.1016/j.neuron.2009.08.012
- Mantini, D., Perrucci, M. G., Del Gratta, C., Romani, G. L., and Corbetta, M. (2007). Electrophysiological signatures of resting state networks in the human brain. *Proc. Natl. Acad. Sci. U.S.A.* 104, 13170–13175. doi: 10.1073/pnas.0700668104
- Mazoyer, B., Zago, L., Mellet, E., Bricogne, S., Etard, O., Houdé, O., et al. (2001). Cortical networks for working memory and executive functions sustain the conscious resting state in man. *Brain Res. Bull.* 54, 287–298. doi: 10.1016/S0361-9230(00)00437-8
- McKhann, G. M. (1984). Clinical diagnosis of Alzheimer's disease. *Neurology* 34:939. doi: 10.1212/WNL.34.7.939
- McKhann, G. M., Knopman, D. S., Chertkow, H., Hyman, B. T., Jack, C. R. Jr, Kawas, C. H., et al. (2011). The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement.* 7, 263–269. doi: 10.1016/j.jalz.2011.03.005
- Mitchell, J. P. (2006). Mentalizing and Marr: an information processing approach to the study of social cognition. *Brain Res.* 1079, 66–75. doi: 10.1016/j.brainres.2005.12.113
- Mo, J., Liu, Y., Huang, H., and Ding, M. (2013). Coupling between visual alpha oscillations and default mode activity. *Neuroimage* 68, 112–118. doi: 10.1016/j.neuroimage.2012.11.058
- Moosmann, M., Ritter, P., Krastel, I., Brink, A., Thees, S., Blankenburg, F., et al. (2003). Correlates of alpha rhythm in functional magnetic resonance imaging

- and near infrared spectroscopy. *Neuroimage* 20, 145–158. doi: 10.1016/S1053-8119(03)00344-6
- Moretti, D. (2004). Individual analysis of EEG frequency and band power in mild Alzheimer's disease. *Clin. Neurophysiol.* 115, 299–308. doi: 10.1016/S1388-2457(03)00345-6
- Moretti, D. V. (2015). Conversion of mild cognitive impairment patients in Alzheimer's disease: prognostic value of Alpha3/Alpha2 electroencephalographic rhythms power ratio. *Alzheimers Res. Ther.* 7, 80. doi: 10.1186/s13195-015-0162-x
- Morris, J. C., Heyman, A., Mohs, R. C., Hughes, J. P., van Belle, G., Fillenbaum, G., et al. (1989). The consortium to establish a registry for Alzheimer's disease (CERAD). Part I. Clinical and neuropsychological assessment of Alzheimer's disease. *Neurology* 39, 1159–1165. doi: 10.1212/WNL.39.9.1159
- Olbrich, S., Mulert, C., Karch, S., Trenner, M., Leicht, G., Pogarell, O., et al. (2009). EEG-vigilance and BOLD effect during simultaneous EEG/fMRI measurement. *Neuroimage* 45, 319–332. doi: 10.1016/j.neuroimage.2008.11.014
- Palva, S., and Palva, J. M. (2007). New vistas for alpha-frequency band oscillations. *Trends Neurosci.* 30, 150–158. doi: 10.1016/j.tins.2007.02.001
- Raichle, M. E., MacLeod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A., and Shulman, G. L. (2001). A default mode of brain function. *Proc. Natl. Acad. Sci. U.S.A.* 98, 676–682. doi: 10.1073/pnas.98.2.676
- Scheeringa, R., Petersson, K. M., Kleinschmidt, A., Jensen, O., and Bastiaansen, M. C. M. (2012). EEG alpha power modulation of fMRI resting-state connectivity. *Brain Connect.* 2, 254–264. doi: 10.1089/brain.2012.0088
- Shirer, W. R., Ryali, S., Rykhlevskaia, E., Menon, V., and Greicius, M. D. (2012). Decoding subject-driven cognitive states with whole-brain connectivity patterns. *Cereb. Cortex* 22, 158–165. doi: 10.1093/cercor/bhr099
- Shulman, G. L., Fiez, J. A., Corbetta, M., Buckner, R. L., Miezin, F. M., Raichle, M. E., et al. (1997). Common blood flow changes across visual tasks: II. Decreases in cerebral cortex. *J. Cogn. Neurosci.* 9, 648–663. doi: 10.1162/jocn.1997.9.5.648
- Spreng, R. N., and Grady, C. L. (2010). Patterns of brain activity supporting autobiographical memory, prospection, and theory of mind, and their relationship to the default mode network. *J. Cogn. Neurosci.* 22, 1112–1123. doi: 10.1162/jocn.2009.21282
- Steininger, S. C., Liu, X., Gietl, A., Wyss, M., Schreiner, S., Gruber, E., et al. (2014). Cortical amyloid beta in cognitively normal elderly adults is associated with decreased network efficiency within the cerebro-cerebellar system. *Front. Aging Neurosci.* 6:52. doi: 10.3389/fnagi.2014.00052
- Tuladhar, A. M., ter Huurne, N., Schoffelen, J.-M., Maris, E., Oostenveld, R., and Jensen, O. (2007). Parieto-occipital sources account for the increase in alpha activity with working memory load. *Hum. Brain Mapp.* 28, 785–792. doi: 10.1002/hbm.20306
- Weiler, M., Teixeira, C. V. L., Nogueira, M. H., Campos, B. M. D., Damasceno, B. P., Cendes, F., et al. (2014). Differences and the relationship in default mode network intrinsic activity and functional connectivity in mild Alzheimer's disease and amnesic mild cognitive impairment. *Brain Connect.* 4, 567–574. doi: 10.1089/brain.2014.0234
- Xia, M., Wang, Z., Dai, Z., Liang, X., Song, H., Shu, N., et al. (2014). Differentially disrupted functional connectivity in posteromedial cortical subregions in Alzheimer's disease. *J. Alzheimers Dis.* 39, 527–543. doi: 10.3233/JAD-131583
- Zhang, H.-Y., Wang, S.-J., Liu, B., Ma, Z.-L., Yang, M., Zhang, Z.-J., et al. (2010). Resting brain connectivity: changes during the progress of Alzheimer disease. *Radiology* 256, 598–606. doi: 10.1148/radiol.10091701
- Zhang, H.-Y., Wang, S.-J., Xing, J., Liu, B., Ma, Z.-L., Yang, M., et al. (2009). Detection of PCC functional connectivity characteristics in resting-state fMRI in mild Alzheimer's disease. *Behav. Brain Res.* 197, 103–108. doi: 10.1016/j.bbr.2008.08.012
- Zheng, W., Liu, X., Song, H., Li, K., and Wang, Z. (2017). Altered functional connectivity of cognitive-related cerebellar subregions in Alzheimer's disease. *Front. Aging Neurosci.* 9:143. doi: 10.3389/fnagi.2017.00143
- Zhou, B., Liu, Y., Zhang, Z., An, N., Yao, H., Wang, P., et al. (2013). Impaired functional connectivity of the thalamus in Alzheimer's disease and mild cognitive impairment: a resting-state fMRI study. *Curr. Alzheimer Res.* 10, 754–766. doi: 10.2174/15672050113109990146
- Zhu, D. C., Majumdar, S., Korolev, I. O., Berger, K. L., and Bozoki, A. C. (2013). Alzheimer's disease and amnesic mild cognitive impairment weaken connections within the default-mode network: a multi-modal imaging study. *J. Alzheimers Dis.* 34, 969–984. doi: 10.3233/JAD-121879
- Zschocke, S., and Hansen, H.-C. (2012). *Klinische Elektroenzephalographie*. Berlin: Springer. doi: 10.1007/978-3-642-19943-1
- Zumer, J. M., Scheeringa, R., Schoffelen, J. M., Norris, D. G., and Jensen, O. (2014). Occipital alpha activity during stimulus processing gates the information flow to object-selective cortex. *PLOS Biol.* 12:e1001965. doi: 10.1371/journal.pbio.1001965

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.

Copyright © 2017 Brueggen, Fiala, Berger, Ochmann, Babiloni and Teipel. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Classifying MCI Subtypes in Community-Dwelling Elderly Using Cross-Sectional and Longitudinal MRI-Based Biomarkers

Hao Guan¹, Tao Liu^{1,2,3*}, Jiyang Jiang^{4,5}, Dacheng Tao^{6,7}, Jicong Zhang^{1,2}, Haijun Niu^{1,3}, Wanlin Zhu^{4,8}, Yilong Wang⁸, Jian Cheng^{9*}, Nicole A. Kochan^{4,5}, Henry Brodaty^{4,10}, Perminder Sachdev^{4,5} and Wei Wen^{4,5}

¹ School of Biological Science and Medical Engineering, Beihang University, Beijing, China, ² Beijing Advanced Innovation Center for Big Data-Based Precision Medicine, Beijing, China, ³ Beijing Advanced Innovation Center for Biomedical Engineering, Beijing, China, ⁴ Centre for Healthy Brain Ageing, School of Psychiatry, University of New South Wales, Sydney, NSW, Australia, ⁵ Neuropsychiatric Institute, Prince of Wales Hospital, Sydney, NSW, Australia, ⁶ UBTech Sydney Artificial Intelligence Institute, Faculty of Engineering and Information Technologies, University of Sydney, Darlington, NSW, Australia, ⁷ The School of Information Technologies, Faculty of Engineering and Information Technologies, University of Sydney, Darlington, NSW, Australia, ⁸ Beijing Tiantan Hospital, Capital Medical University, Beijing, China, ⁹ NIBIB, NICHD, National Institutes of Health, Bethesda, MD, United States, ¹⁰ Dementia Collaborative Research Centre, University of New South Wales, Sydney, NSW, Australia

OPEN ACCESS

Edited by:

Javier Ramirez,
University of Granada, Spain

Reviewed by:

Iman Beheshti,
National Center of Neurology and
Psychiatry, Japan
Heung-Il Suk,
University of North Carolina at Chapel
Hill, United States

*Correspondence:

Tao Liu
tao.liu@buaa.edu.cn
Jian Cheng
jiancheng@ieee.org

Received: 20 July 2017

Accepted: 12 September 2017

Published: 26 September 2017

Citation:

Guan H, Liu T, Jiang J, Tao D, Zhang J, Niu H, Zhu W, Wang Y, Cheng J, Kochan NA, Brodaty H, Sachdev P and Wen W (2017) Classifying MCI Subtypes in Community-Dwelling Elderly Using Cross-Sectional and Longitudinal MRI-Based Biomarkers. *Front. Aging Neurosci.* 9:309. doi: 10.3389/fnagi.2017.00309

Amnesic MCI (aMCI) and non-amnesic MCI (naMCI) are considered to differ in etiology and outcome. Accurately classifying MCI into meaningful subtypes would enable early intervention with targeted treatment. In this study, we employed structural magnetic resonance imaging (MRI) for MCI subtype classification. This was carried out in a sample of 184 community-dwelling individuals (aged 73–85 years). Cortical surface based measurements were computed from longitudinal and cross-sectional scans. By introducing a feature selection algorithm, we identified a set of discriminative features, and further investigated the temporal patterns of these features. A voting classifier was trained and evaluated via 10 iterations of cross-validation. The best classification accuracies achieved were: 77% (naMCI vs. aMCI), 81% (aMCI vs. cognitively normal (CN)) and 70% (naMCI vs. CN). The best results for differentiating aMCI from naMCI were achieved with baseline features. Hippocampus, amygdala and frontal pole were found to be most discriminative for classifying MCI subtypes. Additionally, we observed the dynamics of classification of several MRI biomarkers. Learning the dynamics of atrophy may aid in the development of better biomarkers, as it may track the progression of cognitive impairment.

Keywords: mild cognitive impairment, longitudinal data, early diagnosis, MRI, biomarker, feature selection, machine learning

INTRODUCTION

Mild cognitive impairment (MCI) is thought to be a transitional stage between cognitively normal and dementia (Petersen, 2004). Previous studies have shown that neuroimaging biomarkers are potential predictors of cognitive impairment (Shi et al., 2010; Cuingnet et al., 2011; Davatzikos et al., 2011; Falahati et al., 2014; Trzepacz et al., 2014; Bron et al., 2015; Jung et al., 2016; Lebedeva et al., 2017). Many researchers have developed and implemented machine learning systems which

use neuroimaging biomarkers for more accurate identification of individuals with MCI or dementia (Cui et al., 2012a; Shao et al., 2012; Lebedev et al., 2014; Min et al., 2014; Moradi et al., 2015; Yun et al., 2015; Cai et al., 2017; Guo et al., 2017). Early diagnosis is an essential step in the prevention and early treatment of MCI and dementia.

MCI is clinically heterogeneous with different risks of progression to dementia. Clinical subtypes of MCI have been proposed to broaden the concept, and included prodromal forms of a variety of dementias (Petersen, 2004). MCI is termed “amnesic MCI” (aMCI) when memory loss is the predominant symptom. Almost 10% to 15% aMCI individuals tend to progress to clinically probable Alzheimer’s disease (AD) annually (Grundman et al., 2004). Additionally, MCI is termed “non-amnesic MCI” (naMCI) when impairments are in domains other than memory. Individuals with naMCI were more likely to convert to dementia other than AD, such as vascular dementia or dementia with Lewy bodies (Tabert et al., 2006). The progression of different MCI subtypes to a particular type of dementia has yet to be clearly delineated. On the other hand, MCI does not necessarily lead to dementia, since some studies suggested that MCI subjects have higher rates of reversion to normal cognition than progression to dementia (Brodaty et al., 2013; Pandya et al., 2016). A population-based study found that the reversion rate is lower in aMCI compared with naMCI (Roberts et al., 2014). Reliably identifying MCI of different subtypes would enable more efficient clinical trials and facilitate better targeted treatments.

Longitudinal measurements of Magnetic Resonance Imaging (MRI) in MCI and dementia may provide crucial predictors for tracking the disease progression of dementia (Misra et al., 2009; Risacher et al., 2010; Liu et al., 2013; Mayo et al., 2017). However, only a few studies used longitudinal data for automated classification of MCI and dementia (McEvoy et al., 2011; Li et al., 2012; Zhang et al., 2012a; Ardekani et al., 2017; Huang et al., 2017). Zhang et al. proposed an AD prediction method using longitudinal data which achieved greater classification results than using baseline visit data (Zhang et al., 2012a). Huang et al. presented a longitudinal measurement of MCI brain images and a hierarchical classification method for AD prediction. Their method using longitudinal data consistently outperformed the method using baseline data only (Huang et al., 2017). Despite these efforts, employing machine learning technique with longitudinal MRI features for MCI subtypes classification is rarely studied. And an additional aspect of research when using longitudinal MRI measurements is to identify the biomarkers that remain significant during the time course.

In this study, we used machine learning technique to classify MCI subtypes by employing cross-sectional and longitudinal MRI features. We reported nine independent classification experiments, whereby we compared two groups in each experiment: aMCI vs. cognitively normal (CN), naMCI vs. CN, naMCI vs. aMCI, using features measured at baseline, two-year follow-up, and longitudinally. The longitudinal features were employed by calculating the means and changes of the cross-sectional measurements. Clinical classifications at two-year follow-up were used as the comparison. The features used for classification were cortical surface based, including sulcal width,

cortical thickness, cortical gray matter (GM) volume, subcortical volumes and white matter hyper-intensity (WMH) volume. We compared the classification performance using cross-sectional features and longitudinal features. In addition, we performed feature selection and analyzed the temporal patterns of the selected biomarkers.

MATERIALS AND METHODS

Participants

Participants were members of the Sydney Memory and Aging Study (MAS), a longitudinal study of community-dwelling individuals aged 70–90 years recruited via the electoral roll from two regions of Sydney, Australia (Sachdev et al., 2010). Individuals were excluded at baseline if they had a previous diagnosis of dementia, mental retardation, psychotic disorder including schizophrenia or bipolar disorder, multiple sclerosis, motor neuron disease, developmental disability, or progressive malignancy. The study was approved by the Ethics Committees of the University of New South Wales and the South Eastern Sydney and Illawarra Area Health Service. Written informed consent was obtained from each participant.

Diagnosis

Participants were diagnosed with MCI using the international consensus criteria (Winblad et al., 2004). Specifically, the presence of cognitive impairment as determined by performance on a neuropsychological measure of at least 1.5 standard deviations below published normative values for age and/or education on a test battery covering five cognitive domains (memory, attention/information processing, language, spatial and executive abilities), a subjective complaint of decline in memory or other cognitive function either from the participant or informant, and normal or minimally impaired instrumental activities of daily living attributable to cognitive impairment (total average score <3.0 on the Bayer Activity of Daily Living Scale, Hindmarch et al., 1998).

MCI were classified into two subtypes (aMCI or naMCI) according to cognitive impairment profiles (Petersen, 2004). Participants with no impairments on neuropsychological tests were deemed to have normal cognition. In this study, we included individuals who had MRI scans from both baseline and 2-year follow-up (wave-2), and a wave-2 diagnosis of either cognitively normal or MCI. Demographic characteristics were detailed in **Table 1**. A total of 184 participants met these criteria, including 115 cognitively normal (CN), 42 aMCI, and 27 naMCI. The MRI measurements used in the present study have been previously published (Liu et al., 2013).

Image Acquisition

MRI scans were obtained with a 3-T system (Philips Medical Systems, Best, The Netherlands) using the same sequence for both baseline and follow-up scans: $TR = 6.39$ ms, $TE = 2.9$ ms, flip angle = 8° , matrix size = 256×256 , FOV = $256 \times 256 \times 190$ mm, and slice thickness = 1 mm with no gap, yielding $1 \times 1 \times 1$ mm³ isotropic voxels.

TABLE 1 | Demographic characteristics of the sample.

| Time point | Diagnostic group | No. of subjects (male) | Age mean (SD) | Years of Edu mean (SD) | MMSE score mean (SD) |
|------------|------------------|------------------------|---------------|------------------------|----------------------|
| Baseline | Total | 184 (91) | 77.48 (4.40) | 11.79 (3.60) | 28.16 (1.32) |
| | CN | 117 (56) | 77.12 (4.43) | 11.93 (3.53) | 28.39 (1.23) |
| | aMCI | 40 (28) | 78.36 (4.11) | 11.81 (3.96) | 27.58 (1.32) |
| | naMCI | 27 (7) | 77.76 (4.65) | 11.14 (3.42) | 28.00 (1.44) |
| Wave-2 | Total | 184 (91) | 79.38 (4.40) | 11.79 (3.60) | 28.40 (1.41) |
| | CN | 115 (53) | 78.78 (4.15) | 12.06 (3.42) | 28.83 (1.16) |
| | aMCI | 42 (30) | 81.26 (4.98) | 11.87 (4.16) | 27.64 (1.59) |
| | naMCI | 27 (8) | 79.03 (3.72) | 10.49 (3.27) | 27.78 (1.40) |

CN, cognitively normal; aMCI, amnesic mild cognitive impairment (MCI); naMCI, non-amnesic MCI; Edu, education, MMSE, Mini-mental state examination.

Image Processing

Sulcal Measures

Cortical sulci were extracted from the images via the following steps. First, non-brain tissues were removed to produce images containing only GM, white matter (WM) and cerebrospinal fluid (CSF). This was done by warping a brain mask defined in the standard space back to the T1-weighted structural MRI scan. The brain mask was obtained with an automated skull stripping procedure based on the SPM5 skull-cleanup tool (Ashburner, 2009). Individual sulci were identified and extracted using the BrainVisa (BV, version 3.2) sulcal identification pipeline (Rivière et al., 2009). A sulcal labeling tool incorporating 500 artificial neural network-based pattern classifiers (Rivière et al., 2002; Sun et al., 2007) was used to label sulci. Sulci that were mislabeled by BV were manually corrected. For each hemisphere, we determined the average sulcal width for five sulci: superior frontal, intra-parietal, superior temporal, central, and the sylvian fissure. Sulcal width was defined as the average 3D distance between opposing gyral banks along the normal projections to the medial sulcal mesh (Kochunov et al., 2012). The five sulci investigated in the present study were chosen because they were present in all individuals, large and relatively easy to identify after facilitating error detection and correction, and located on different cerebral lobes. For each hemisphere, we calculated the global sulcal index (g-SI) as the ratio between the total sulcal area and outer cortical area (Penttilä et al., 2009). We calculated the g-SI of each brain with no manual intervention using BV.

Cortical Thickness, GM Volume

We computed average regional GM volume, average regional cortical thickness using the longitudinal stream in FreeSurfer 5.1 (<http://surfer.nmr.mgh.harvard.edu/>) (Reuter et al., 2012). This stream specifically creates an unbiased specific within-subject template space and image using robust, inverse consistent registration (Reuter and Fischl, 2011; Reuter et al., 2012). Briefly, this pipeline included the following processing steps, skull stripping, Talairach transforms, atlas registration, spherical surface maps, and parcellation of cerebral cortex (Desikan et al., 2006; Reuter et al., 2012). We applied Desikan parcellation (Desikan et al., 2006) which resulted 34 cortical regions of interest (ROIs) in each hemisphere. We visually inspected registration

and segmentation. Scans were excluded if they failed visual quality control, resulting in an unequal number of scans available for different brain structures. We calculated both the cortical thickness and the regional volumes for every cortical regions of the Desikan parcellation.

Subcortical Volume

Subcortical brain structures were extracted using FSL's FIRST (FMRIB Image Registration and Segmentation Tool, Version 1.2), a model-based segmentation/registration tool (Patenaude et al., 2011). We included the following left and right subcortical structures: thalamus, caudate, putamen, pallidum, hippocampus, amygdala, and nucleus accumbens. Briefly, the FIRST algorithm modeled each participant's subcortical structure as a surface mesh, using a Bayesian model incorporating a training set of all images. We conducted visual quality control of FSL results using ENIGMA protocols (<http://enigma.ini.usc.edu/>). Three slices of each of coronal, sagittal and axial planes were extracted from each linearly transformed brain. For comparison, an outline of the templates was mapped onto the slices. We confirmed that the size of the participant brain corresponded with that of the template, verified that the lobes were appropriately situated, and confirmed that the orientation of the participant matched the template.

WMHs

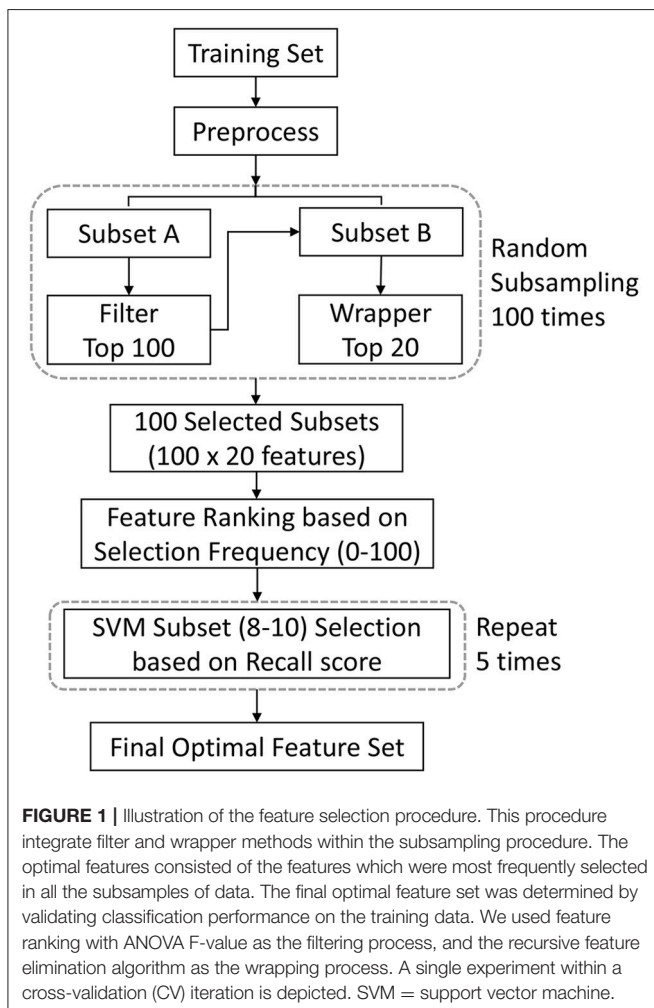
WMHs were delineated from coronal plane 3D T1-weighted and Fluid Attenuated Inversion Recovery (FLAIR) structural image scans using a pipeline described in detail previously (Wen et al., 2009). For each hemisphere, we calculated WMH volumes of eight brain regions: temporal, frontal, occipital, parietal, ventricle body, anterior horn, posterior horn, and cerebellum.

We obtained neuroimaging measurements of all participants at baseline and wave-2. The changes and the means values of those measurements were considered as the longitudinal features. There were altogether 178 MRI measurements for baseline and wave-2 feature sets, which included 12 sulcal measurements, 68 thickness measurements, 68 volume measurements, 14 subcortical measurements, and 16 WMH measurements. With the means and the changes, the longitudinal feature set included 356 MRI measurements.

Feature Selection

The aims of feature selection were to maximize the performance of classification by identifying the most discriminative features, and help in understanding the neuropathological basis of neurocognitive impairments such as MCI and dementia. Supervised feature selection methods were often divided into three categories, namely “filter,” “wrapper,” and “embedded,” respectively (Mwangi et al., 2014). A particular problem of those methods was that when they were applied in the neuroimaging fields, where the number of features largely exceeded the number of examples, the cross-validation based error estimates usually led to results with extremely large variances (Dougherty et al., 2010; Tohka et al., 2016). We proposed a feature selection method in this study to reduce the variances by integrating the filter and the wrapper procedures within the subsampling iterations. The optimal feature subset consisted of the features which were most frequently selected in all the subsamples of data. The discriminative abilities of the features were assessed in terms of the selection frequencies.

Figure 1 shows the flowchart of the feature selection procedure used in our study. We first randomly subsampled the training set 100 times. During each subsampling iteration,

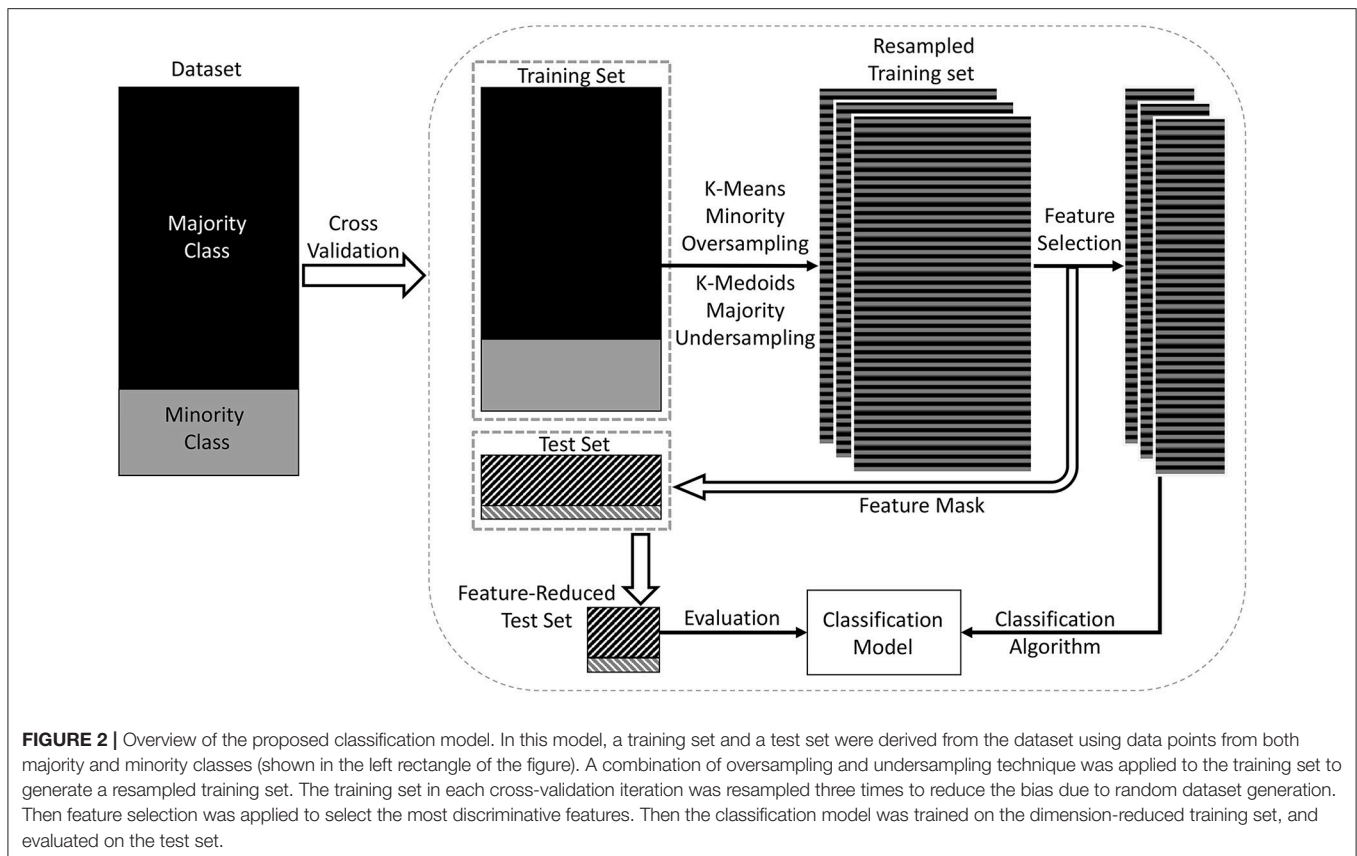


data were divided into two subsets of equal size, subset A and subset B. Subset A was processed by a filter to select features. The selected features were then applied to subset B. The subset B was processed by a wrapper to further reduce the number of features. After the subsampling processes, features were subsequently ranked in order of selection frequencies. The final optimal feature set was then determined by validating classification performance on the training data, using features chosen on the basis of frequency rank thresholds.

In the filter stage, ANOVA (analysis of variance) *F*-value were used to rank features on the basis of correlations with their diagnostic label. The top 100 features were selected at this stage. Then in the wrapping stage, the recursive feature elimination algorithm (Guyon et al., 2002) was used to further remove less informative features. Among the top 100 features, 20 were retained in this stage. The selection frequencies could be 100 at maximum or 0 at minimum. To mitigate the curse-of-dimensionality problem, the final feature set was limited with less than 10 features, and a variation section was established for the feature set to achieve the best validation performance. Given a frequency rank threshold N_f ($N_f \in [10, 9, 8]$), we randomly split the training data into 2 subgroups: one for training a SVM (Vapnik, 1995) classifier with top N_f features, and the other for validation. The kernel for the SVM is the radial basic function (rbf). This step was repeated 5 times, and the recall scores were computed (the recall score is the ratio $Tp/(Tp + Fn)$, where Tp is the number of true positives and Fn is the number of false negatives). We chose the recall score as the criteria to minimize the impact of sample proportion imbalance. The top N_f features with the highest average recall score became the optimal feature set. We also evaluated the selected features using 2-tailed *t*-test.

Classification and Validation

The imbalance of the sample could lead to a suboptimal classification performance. This study investigated a population-based sample, consisting of more cognitively normal individuals than MCI. There was also a large difference between the sample sizes of different MCI subtypes. We addressed this problem by using the data-resampling technique (Chawla et al., 2002; Dubey et al., 2014). An overview of the procedure is shown in **Figure 2**. We used a combination of oversampling and undersampling (Batista et al., 2004). K-means clustering (Macqueen, 1967) algorithm was used for oversampling, where new synthetic data were generated by clustering the minority class data. Briefly, N_s samples were clustered into $N_s/3$ clusters, and $N_s/3$ centroids were generated. Then these centroids and the original samples were combined for the next iteration of oversampling. The oversampling procedure was repeated until the size of minority class was $2/3$ the size of the majority class. K-Medoids clustering (Hastie et al., 2001) algorithm was used for undersampling, where actual data points from the majority class were chosen as the cluster centers. The final training set was a combination of the oversampled minority class data and the undersampled majority class data. While resampling the training set, the test set remained the same. The training set was resampled 3 times to reduce the bias due to random data generation. Then the



feature selection method was applied on those resampled training sets, thus producing 3 learning models. These models were combined using majority voting, where the final label of an instance was decided based on the majority votes received from all the models.

We chose Voting Classifier for classification (Maclin and Opitz, 1999). A Voting Classifier combines conceptually different machine learning classifiers and uses a majority vote or the average predicted probabilities (soft vote) to predict the class labels. The advantage of Voting Classifier is to balance out the individual weaknesses of a set of equally well performing models. We chose SVM (rbf kernel), Logistic Regression (LR) (Cox, 1958), and Random Forest (RF) (Breiman, 2001) as the estimators of the Voting Classifier. All the estimators were with default settings of parameters. Specific weights (1:4:1) were assigned to SVM, LR and RF via the weights parameter. The weights were selected experimentally to aim at a better sensitivity score. We started with the equal weights (1:1:1), and changed the weights to obtain the best results. The predicted class probabilities of each classifier were collected, multiplied by the weights of classifiers, and averaged. The final class label was then derived from the class label with the highest average probability. As different features had different scales, we standardized all the training data within a 0–1 range, and the same procedure was then applied to the test data.

We evaluated our method using stratified Shuffle Split cross-validation procedure, also known as Monte Carlo cross-validation (Berrar et al., 2007), which returned stratified randomized folds by preserving the percentage of samples for each class. The cross-validation procedure was repeated 10 times with a fixed 9:1 train-test ratio. The final classification results represented the average of these 10 independent experiments. We applied four metrics to assess the performance of the model: the accuracy, the specificity, the sensitivity, and the area under the receiver operating characteristic curve (AUC). AUC is a better measure than accuracy in imbalanced data sets and real-world applications (Huang and Ling, 2005; Bekkar et al., 2013).

It was important to note that we obtained a unique set of selected features in each training set. The training set in each cross-validation iteration was resampled 3 times, thus producing 3 resampled training sets. In each training set, the maximum possible selection frequency of one feature was 100. Considering the feature selection and data-resampling steps within the 10-iteration cross-validation procedure, the final maximum possible selection frequency of each feature was $3 \times 100 \times 10 = 3,000$.

All the data processing and analyzing were performed using Python libraries Numpy 1.10.4 (Walt et al., 2011) and Scipy 0.17.0 (Jones et al., 2001) on Python 2.7.11 (Anaconda 4.0.0–64 bit, <http://www.continuum.io/>). All the machine learning methods were performed using the library Scikit-Learn 0.17.1 (Pedregosa et al., 2011).

RESULTS

MCI Subtypes Classification

As shown in **Table 2**, in the classification of aMCI and CN, compared with using baseline features, using longitudinal features improved the performance to accuracy of 73%, sensitivity of 53%, specificity of 80%, and AUC of 0.75; the results of using longitudinal features were not superior to that using wave-2 features. Identifying naMCI from CN was relatively difficult considering the poor sensitivity value and AUC; the results of using longitudinal and cross-sectional features were comparable and without significant difference. In the classification of naMCI vs. aMCI, compared with using longitudinal features, using baseline features achieved better performance; the results of using wave-2 features were not significantly different from using longitudinal features.

Discriminative Features

The discriminative ability of the features used in this study were assessed by examining the frequency with which they were selected. We listed the top 10 most frequently selected features in each MCI subtype classification experiment (see **Tables 3–5**). In the comparison of aMCI vs. CN, thickness of right frontal pole, left superior temporal, volume of right thalamus, and right hippocampus were more discriminative than the rest of features (see **Table 3**). In the classification of naMCI vs. aMCI, thickness of right rostral middle frontal, right pericalcarine, right frontal pole, and volume of right rostral anterior cingulate were more discriminative than the others (see **Table 5**). Regardless of cross-sectional (baseline and wave-2) or longitudinal, all the features mentioned above were listed in the top-10 feature list. In the naMCI vs. CN comparison, volume of left temporal pole and right amygdala were also discriminative (see **Table 4**).

The top-10 selected features were analyzed to identify the temporal patterns. Several features measured at different time points showed dynamic discriminative powers. **Figures 3–5** shows the selection frequencies of the stable features measured at each time point. A feature may be identified as stable when this

feature was selected at all the baseline, wave-2, and longitudinally. The selection frequencies of the stable features for aMCI vs. CN classification are shown in **Figure 3**. We observed that thickness of right frontal pole was a stable biomarker, since its selection frequencies were close between different time points. The selection frequencies of several biomarkers changed visibly over time, including volume of right thalamus, right hippocampus, and thickness of left superior temporal. In the classification of naMCI vs. CN (see **Figure 4**), only a few features were stable. We observed that the volume of right amygdala provided more useful information at baseline. Volume of left temporal pole and right rostral cingulate carried more information at baseline. In the classification of naMCI vs. aMCI (see **Figure 5**), volume of right rostral middle frontal and thickness of right pericalcarine thickness were selected more often at baseline, while volume of right frontal pole were more discriminative at wave-2. And volume of right rostral anterior cingulate provided important information at all-time points.

Furthermore, some features were selected in the top-10 feature list at either baseline or wave-2, such as the right g-SI index, sulcal width of superior frontal (see **Table 3**); thickness of left lateral occipital, WMH volume of right cerebellum (see **Table 4**); thickness of right lateral occipital, and WMH volume of right frontal (see **Table 5**). On the other hand, some features were selected only in longitudinal cases, such as sulcal width of right superior temporal, thickness of left inferior temporal (see **Table 3**); volume of right entorhinal and right posterior cingulate, thickness of left posterior cingulate and temporal pole (see **Table 4**); thickness of left precentral, volume of right entorhinal (see **Table 5**). Most of these longitudinal features were the differences (changes value) between the measures of two time points.

DISCUSSION

Our study examined classification of MCI subtypes in community-dwelling elderly using cross-sectional and

TABLE 2 | Classification results of MCI subtypes: features measured at baseline, wave-2 and longitudinally are used and compared.

| Task | No. of minority class | No. of majority class | Method | Accuracy (%) | Sensitivity (%) | Specificity (%) | AUC |
|----------------|-----------------------|-----------------------|--------------|--------------|-----------------|-----------------|-------|
| aMCI vs. CN | aMCI = 42 | CN = 115 | Baseline | 0.64 | 0.42 | 0.71 | 0.68 |
| | | | Wave-2 | 0.81* | 0.68 | 0.85 | 0.74 |
| | | | Longitudinal | 0.73 | 0.53 | 0.80 | 0.75 |
| naMCI vs. CN | naMCI = 27 | CN = 115 | Baseline | 0.67 | 0.37 | 0.75 | 0.57 |
| | | | Wave-2 | 0.65 | 0.30 | 0.74 | 0.58 |
| | | | Longitudinal | 0.70 | 0.23 | 0.82 | 0.60 |
| naMCI vs. aMCI | naMCI = 27 | aMCI = 42 | Baseline | 0.77* | 0.70* | 0.82 | 0.84* |
| | | | Wave-2 | 0.71 | 0.57 | 0.82 | 0.70 |
| | | | Longitudinal | 0.61 | 0.40 | 0.78 | 0.71 |

wave-2, 2-year follow-up; MCI, mild cognitive impairment; CN cognitively normal; aMCI, amnesic MCI; naMCI, non-amnesic MCI; AUC, area under the receiver operating characteristic curve.

*Significantly different from the method using longitudinal features; results are from *t*-test ($p < 0.05$).

TABLE 3 | Selected features for the classification of aMCI vs. CN.

| Task | Baseline feature | Frequency | p ^a | Wave-2 feature | Frequency | p ^a | Longitudinal feature | Frequency | p ^a |
|-------------|--|--------------|----------------|---|--------------|------------------|---|--------------|------------------|
| aMCI vs. CN | Right frontal pole thickness | 2,658 | 0.001 | Right frontal pole thickness | 2,786 | <0.001 | Right frontal pole thickness | 2,715 | <0.001 |
| | Right thalamus volume | 2,136 | 0.006 | Right hippocampus volume | 2,105 | <0.001 | Right thalamus volume | 2,470 | 0.001 |
| | Left superior temporal thickness | 1,620 | 0.001 | Right thalamus volume | 1,775 | 0.001 | Right hippocampus volume | 2,071 | 0.001 |
| | Right hippocampus volume | 1,344 | 0.005 | Left superior temporal thickness | 1,144 | 0.005 | Left superior temporal thickness | 1,212 | 0.002 |
| | Right g-SI ^c | 1,265 | 0.003 | Left sulcal width of superior frontal | 1,062 | 0.002 | Right sulcal width of superior temporal ^{b*} | 1,036 | 0.027 |
| | Right transverse temporal thickness ^c | 755 | 0.013 | Right sulcal width of superior frontal ^c | 963 | 0.001 | Right pericalcarine thickness | 876 | 0.004 |
| | Right pericalcarine thickness | 736 | 0.005 | Right amygdala volume ^c | 963 | 0.073 | Left precentral thickness ^{b*} | 869 | 0.035 |
| | Right rostral anterior cingulate volume ^c | 693 | 0.128 | Right pericalcarine thickness | 958 | 0.006 | Left inferior temporal thickness [*] | 827 | 0.019 |
| | Right paracentral thickness ^c | 637 | 0.022 | Right accumbens volume ^c | 660 | 0.011 | Right paracentral thickness ^{b*} | 819 | 0.012 |
| | Left posterior cingulate volume ^c | 545 | 0.134 | Left medial orbitofrontal thickness ^c | 633 | 0.045 | Right sulcal width of superior frontal | 722 | 0.002 |

A feature measured at baseline, wave-2 or longitudinally is defined as baseline feature, wave-2 feature or longitudinal feature, respectively. The first 10 most frequently selected features and their selection frequencies are listed. The maximum possible selection frequency of each feature is 3000. The features with selection frequencies above 1500 are in bold. wave-2, 2-year follow-up; MCI, mild cognitive impairment; CN, cognitively normal; aMCI, amnesic MCI.

^aResults for comparisons of positive subjects and negative subjects using t-tests.

^bChanges measurements, the rest longitudinal features are means measurements.

^cFeatures that were selected at a single time point (either at baseline or wave-2).

^{*}Features that were selected only in longitudinal case.

longitudinal MRI measurements. Our classification framework implemented a data-resampling step to reduce the effect of the class-imbalance, and a feature selection step in which maximally most discriminative feature subsets were identified. The results suggested that individuals with aMCI could be differentiated from CN and naMCI with MRI-based biomarkers, but identifying naMCI from CN was still a challenge. Identifying aMCI from CN using longitudinal features achieved better performance than that using baseline features, but the results were not superior to that using wave-2 features. The best performance of differentiating aMCI from naMCI was achieved with baseline features. In addition, we analyzed and identified the dynamics of the biomarkers.

The subtlety of brain changes in MCI challenges the image-based classification. Previous studies reported using machine learning to differentiate MCI from cognitively normal (Wee et al., 2011, 2012; Zhang et al., 2011, 2017; Cui et al., 2012b; Liu et al., 2015, 2017). Cui et al. used combined measurements of T1-weighted and diffusion tensor imaging (DTI) to distinguish aMCI from CN, achieved a classification accuracy of 71%, sensitivity 52%, specificity 78%, and AUC 0.70 (Cui et al., 2012b). Our performance (accuracy 81%, sensitivity 68%, specificity 85%, and AUC 0.74) is better than their study. The approach of Wee et al. was a kernel combination method that utilized DTI and resting-state functional magnetic resonance imaging (Wee et al., 2012). Although their classification accuracy of 96.3% is higher than ours, the inclusion of multi-modality imaging could

restrict their use in clinical settings, and the small sample size of fewer than 30 participants may also make their results less robust. Considering the heterogeneity of MCI, we performed MCI subtypes classification, and the results demonstrated that aMCI and naMCI could be accurately separated with MRI biomarkers. And the results showed that the various groups demonstrated different patterns of atrophy on MRI. However, differentiating naMCI from CN was difficult considering the low sensitivities (see Table 2). The serious imbalance of classes could result in this poor performance, although we had performed data-resampling to mitigate the difference of the sample sizes. Compared with aMCI, naMCI individuals are more likely to revert to normal cognition (Roberts et al., 2014; Aerts et al., 2017). The MCI individuals who reverted might have different underlying mechanisms (Zhang et al., 2012b). In addition, higher estimates of MCI incidence in clinic-based studies (Petersen, 2004, 2010) than in population-based studies suggested that the rate of reversion to normal cognition may be lower in the clinic setting than in population-based studies (Koepsell and Monsell, 2011; Lopez et al., 2012) such as ours.

Longitudinal patterns of atrophy identified in MRI measurements can be used to elevate the prediction of cognitive decline (Rusinek et al., 2003; Risacher et al., 2010). McEvoy et al. investigated whether single-time-point and longitudinal volumetric MRI measures provided predictive prognostic information in patients with aMCI. Their results showed that the information regarding the rate of atrophy progression

TABLE 4 | Selected features for the classification of naMCI vs. CN.

| Task | Baseline feature | Frequency | p ^a | Wave-2 feature | Frequency | p ^a | Longitudinal feature | Frequency | p ^a |
|--------------|---|--------------|----------------|---|--------------|----------------|--|--------------|------------------|
| naMCI vs. CN | Right WMH volume of cerebellum^c | 2,240 | 0.014 | Left lateral occipital thickness^c | 2,087 | 0.002 | Right entorhinal volume^{b*} | 2,866 | <0.001 |
| | Left temporal pole volume | 2,227 | 0.072 | Right rostral middle frontal thickness | 1,670 | 0.024 | Right amygdala volume | 1,852 | 0.001 |
| | Right amygdala volume | 2,027 | 0.002 | Left temporal pole volume | 1,636 | 0.086 | Right posterior cingulate volume^{b*} | 1,608 | 0.008 |
| | Right rostral middle frontal thickness | 1,757 | 0.008 | Right amygdala volume | 1,527 | 0.003 | Left lateral occipital thickness | 1,434 | 0.002 |
| | Right rostral anterior cingulate volume | 1,718 | 0.011 | Right sulcal width of superior frontal ^c | 1,259 | 0.017 | Left temporal pole volume | 1,256 | 0.074 |
| | Left middle temporal thickness ^c | 1,316 | 0.002 | Left pericalcarine volume ^c | 1,218 | 0.012 | Left posterior cingulate thickness ^{b*} | 891 | 0.022 |
| | Right inferior parietal thickness ^c | 953 | 0.002 | Right rostral anterior cingulate volume | 993 | 0.026 | Left amygdala volume ^{b*} | 746 | 0.117 |
| | Right thalamus volume ^c | 833 | 0.005 | Right putamen volume ^c | 754 | 0.001 | Left temporal pole thickness ^{b*} | 630 | 0.054 |
| | left transverse temporal volume ^c | 778 | 0.182 | Right supramarginal volume | 602 | 0.263 | Left middle temporal thickness | 613 | 0.007 |
| | Right supramarginal volume | 638 | 0.108 | Left sulcal width of superior temporal ^c | 515 | 0.134 | Right WMH volume of cerebellum | 578 | 0.378 |

A feature measured at baseline, wave-2 or longitudinally is defined as baseline feature, wave-2 feature or longitudinal feature, respectively. The first 10 most frequently selected features and their selection frequencies are listed. The maximum possible selection frequency of each feature is 3000. The features with selection frequencies above 1,500 are in bold. wave-2, 2-year follow-up; CN, cognitively normal; naMCI, non-amnesic MCI.

^aResults for comparisons of positive subjects and negative subjects using t-tests.

^bChanges measurements, the rest longitudinal features are means measurements.

^cFeatures that were selected at a single time point (either at baseline or wave-2).

*Features that were selected only in longitudinal case.

over a 1-year period improved risk prediction compared with using single-time-point MRI measurement (McEvoy et al., 2011). Huang et al. used longitudinal changes over 4 years of T1-weighted MRI scans to predict AD conversion in MCI subjects. Their results showed that the model with longitudinal data consistently outperformed the model with baseline data, especially achieved 17% higher sensitivity than the model with baseline data (Huang et al., 2017). In our study, the results showed that the longitudinal features failed to provide additional information for identifying aMCI and naMCI compared with cross-sectional features. In the classification of aMCI vs. CN, the accuracy with longitudinal features was nearly 10% higher than the accuracy with baseline features, but was not superior to the accuracy with wave-2 features (Table 2). The performance of using longitudinal features was comparable to using cross-sectional features at baseline and wave-2 for distinguishing naMCI from CN. In addition, the highest performance of distinguishing naMCI from aMCI was achieved with baseline features (see Table 2). This might be because the progression of naMCI showed no coherent pattern of atrophy. The patterns of atrophy differ among aMCI and naMCI, and subjects with naMCI showed scattered patterns of gray matter loss without any particular focus (Whitwell et al., 2007). All the subjects of our study were community-dwelling. It was likely that the naMCI subjects had atrophy patterns closer to those of CN at baseline, but over the time the patterns progressed to more MCI-like at wave-2. Our results also indicated that features selected for identifying naMCI were unstable over time, which

might be because clinical classification of naMCI can be based on impairment individually or in combination across a range of non-amnesic cognitive domains (language, visuo-spatial, processing speed, or executive abilities).

Longitudinal research has observed the dynamics of biomarkers (Trojanowski et al., 2010; Sabuncu et al., 2011; Eskildsen et al., 2013; Zhou et al., 2013). Some features provided significant information at all-time points while some other features were shown to be useful at a specific time point. Eskildsen et al. demonstrated that prediction accuracies of conversion from MCI to AD can be improved by learning the atrophy patterns that were specific to the different stages of disease progression (Eskildsen et al., 2013). They found that medial temporal lobe structures were stable biomarkers across all stages. Hippocampus was not discriminative at 36 months prior to AD diagnosis, but was included in all prediction cases of later stages. In addition, biomarkers were mostly selected from the cingulate gyrus, which is well known to be affected in early AD (Eskildsen et al., 2013). Histological studies suggest that the integrity of entorhinal cortex is among the first affected, which is then only later followed by an atrophy of the hippocampus (Braak et al., 1993). In our study, we also found that volume of the right hippocampus was more discriminative at wave-2 (see Figure 3, Table 3), which would complemented the histological findings. Furthermore, the thalamic volume was discriminative and stable over time (see Figure 3, Table 3), which was consistent with a previous study that the structure and function of thalamus determined severity of cognitive impairment (Schoonheim

TABLE 5 | Selected features for the classification of naMCI vs. aMCI.

| Task | Baseline feature | Frequency | p ^a | Wave-2 feature | Frequency | p ^a | Longitudinal feature | Frequency | p ^a |
|----------------|--|--------------|------------------|--|--------------|------------------|--|--------------|------------------|
| naMCI vs. aMCI | Right rostral middle frontal thickness | 2,643 | <0.001 | Right rostral anterior cingulate volume | 2,754 | <0.001 | Right rostral anterior cingulate volume | 2,598 | <0.001 |
| | Right rostral anterior cingulate volume | 2,538 | <0.001 | Right frontal pole thickness | 2,634 | <0.001 | Right rostral middle frontal thickness | 2,502 | <0.001 |
| | Right pericalcarine thickness | 2,241 | 0.001 | Right rostral middle frontal thickness | 2,190 | <0.001 | Right frontal pole thickness | 2,478 | <0.001 |
| | Right frontal pole thickness | 1,815 | <0.001 | Right pericalcarine thickness | 1,551 | 0.005 | Right pericalcarine thickness | 1,830 | 0.002 |
| | Right g-si^c | 1,539 | 0.004 | Left transverse temporal volume ^c | 1,131 | 0.028 | Right lateral occipital thickness | 1,062 | 0.001 |
| | Right lateral occipital thickness ^c | 1,023 | <0.001 | Right wmh volume of frontal ^c | 1,071 | 0.122 | Right entorhinal volume ^{d*} | 1,029 | 0.010 |
| | Right transverse temporal thickness ^c | 750 | 0.014 | Left rostral middle frontal volume ^c | 813 | 0.037 | Left transverse temporal thickness | 867 | 0.009 |
| | Left inferior temporal thickness ^c | 687 | <0.001 | Right insula thickness ^c | 678 | 0.037 | Left inferior temporal thickness | 666 | 0.001 |
| | Right parsorbitalis thickness ^c | 666 | 0.002 | Right frontal pole volume ^c | 573 | 0.054 | Left precentral thickness* | 639 | 0.001 |
| | Right transverse temporal thickness ^c | 480 | 0.011 | Right sulcal width of superior temporal ^c | 552 | 0.026 | Right g-SI | 591 | 0.011 |

A feature measured at baseline, wave-2 or longitudinally is defined as baseline feature, wave-2 feature or longitudinal feature, respectively.

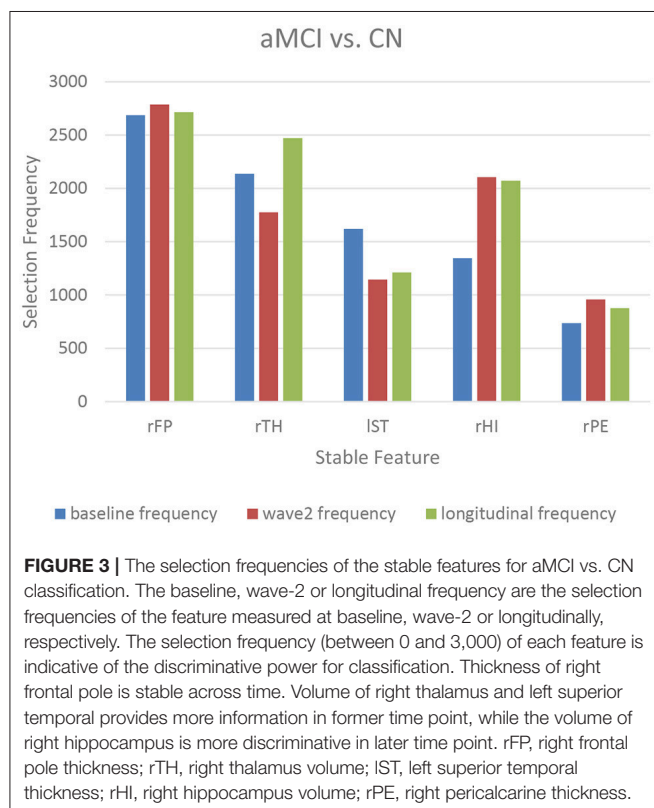
The first 10 most frequently selected features and their selection frequencies are listed. The maximum possible selection frequency of each feature is 3,000. The features with selection frequencies above 1500 are in bold. Key: wave-2, 2-year follow-up; aMCI, amnesic MCI; naMCI, non-amnesic MCI.

^aResults for comparisons of positive subjects and negative subjects using t-tests.

^bChange measurement, the rest longitudinal features are mean measurements.

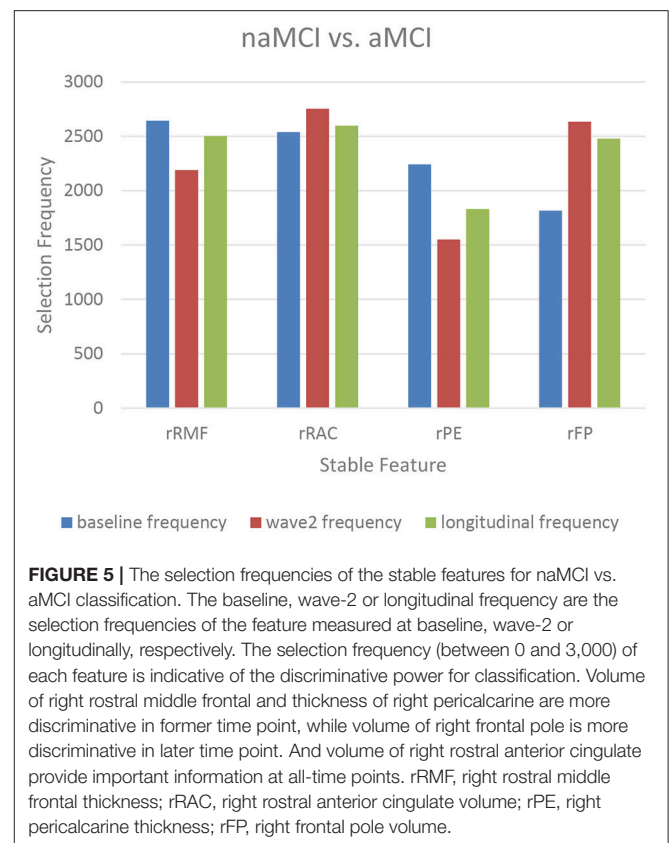
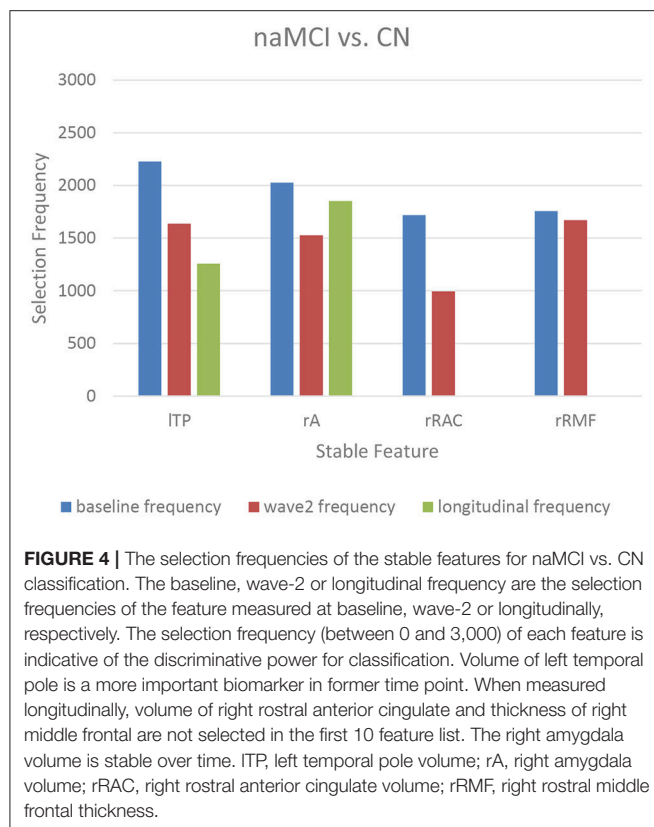
^cFeatures that were selected at a single time point (either at baseline or wave-2).

*Features that were selected only in longitudinal case.



et al., 2015). Volume of left posterior cingulate and right rostral anterior cingulate were more discriminative at baseline for identifying aMCI and naMCI from CN (see **Tables 3, 4**), while volume of right rostral anterior cingulate was a stable biomarker for naMCI vs. aMCI classification over time (see **Figure 5, Table 5**). Zhou et al. used the baseline MRI features to predict MMSE (The Mini-Mental State Examination, Folstein et al., 1975) and ADAS-Cog (Alzheimer's Disease Assessment Scale cognitive subscale, Rosen et al., 1984) scores in the next 4 years (Zhou et al., 2013). They observed that the average cortical thickness of left middle temporal, left and right entorhinal, and volume of left hippocampus were important biomarkers for predicting ADAS-Cog scores at all-time points. Cortical volume of left entorhinal provided significant information in later stages than in the first 6 months. Several biomarkers including volume of left and right amygdala provided useful information only at later time points (Zhou et al., 2013). In our study, cross-sectional (both baseline and wave-2) volume of right entorhinal was not an important biomarker for the classification of naMCI vs. CN, but the longitudinal volume change of right entorhinal (see **Table 4**) was discriminative. Volume of right amygdala was discriminative at all-time points for naMCI vs. CN classification (see **Figure 4, Table 4**). The dynamics of biomarker could potentially aid in developing stable imaging biomarkers and in tracking the progression of cognitive impairment.

The use of same dataset for feature selection and classification is termed "double-dipping," which will lead to distorted



descriptive statistics and artificially inflated accuracies (Kriegeskorte et al., 2009; Pereira et al., 2009; Eskildsen et al., 2013; Mwangi et al., 2014). Due to the limited samples in neuroimaging studies, carelessly designed training, testing and validation schemes, the risk of double-dipping is high. Eskildsen et al. used cortical regions potentially discriminative for predicting AD. They found that by inclusion of test subjects in the feature selection process, the prediction accuracies were artificially inflated (Eskildsen et al., 2013). In our experiments, training datasets and test datasets were adequately separated using cross-validation procedure. The training set in each cross-validation iteration were used for data-resampling, feature selection and classifier training, while the test set were only used for validating classification performance.

The main limitation of the present study was the limited sample size. Our method required longitudinal data, thus limiting the subjects with MRI scans at both time points. Secondly, this study investigated a population-based sample, consisting of more cognitively normal individuals than MCI. There was also a difference between the sample sizes of aMCI and naMCI. The findings need to be replicated in other data sets.

CONCLUSION

In conclusion, the present study investigated MCI subtypes classification in a sample from community-dwelling elderly using both cross-sectional and longitudinal MRI features.

Our experiments suggested that longitudinal features were not superior to the cross-sectional features for MCI subtypes classifications. Dynamics of the biomarkers were analyzed and identified. Future studies with longer follow-up and more measurement occasions may lead to the better understanding of the trajectories for cognitive impairment.

AUTHOR CONTRIBUTIONS

HG, TL, and JC: Study design, data analyses, interpretation of the results, manuscript writing. WW and DT: Study design, interpretation of the results. NK, PS, and HB: Data collection, interpretation of the results. JJ, JZ, HN, WZ, and YW: Data analyses. All authors participated in manuscript revision and final approval.

FUNDING

This research received support from the Natural Science Foundation of China [grant numbers 81401476], the National Key Research and Development Program of China [grant numbers 2016YFF0201002], the National Health and Medical Research Council (NHMRC) Program Grants [grant numbers 350833, 56896, 109308], and the Australian Research Council Projects [grant numbers FL-170100117, DP-140102164, LP-150100671].

REFERENCES

- Aerts, L., Heffernan, M., Kochan, N. A., Crawford, J. D., Draper, B., Trollor, J. N., et al. (2017). Effects of MCI subtype and reversion on progression to dementia in a community sample. *Neurology* 88, 2225–2232. doi: 10.1212/WNL.0000000000004015
- Ardekani, B. A., Bermudez, E., Mubeen, A. M., Bachman, A. H., and Alzheimer's Disease Neuroimaging Initiative (2017). Prediction of incipient alzheimer's disease dementia in patients with mild cognitive impairment. *J. Alzheimers. Dis.* 55, 269–281. doi: 10.3233/JAD-160594
- Ashburner, J. (2009). Computational anatomy with the SPM software. *Magn. Reson. Imaging* 27, 1163–1174. doi: 10.1016/j.mri.2009.01.006
- Batista, G. E. A. P. A., Prati, R. C., and Monard, M. C. (2004). A study of the behavior of several methods for balancing machine learning training data. *SIGKDD Explor. Newsl.* 6, 20–29. doi: 10.1145/1007730.1007735
- Bekkar, M., Djemaa, H. K., and Alitouch, T. A. (2013). Evaluation measures for models assessment over imbalanced data sets. *J. Inform. Eng. Appl.* 3, 27–38.
- Berrard, D., Granzow, M., and Dubitzky, W. (2007). "Introduction to genomic and proteomic data analysis," in *Fundamentals of Data Mining in Genomics and Proteomics* (Boston, MA: Springer).
- Braak, H., Braak, E., and Bohl, J. (1993). Staging of Alzheimer-related cortical destruction. *Eur. Neurol.* 33, 403–408. doi: 10.1159/000116984
- Breiman, L. (2001). Machine learning, volume 45, Number 1 - SpringerLink. *Mach. Learn.* 45, 5–32. doi: 10.1023/A:1010933404324
- Brodsky, H., Heffernan, M., Kochan, N. A., Draper, B., Trollor, J. N., Reppermund, S., et al. (2013). Mild cognitive impairment in a community sample: the sydney memory and ageing study. *Alzheimers Dement.* 9, 310–317.e311. doi: 10.1016/j.jalz.2011.11.010
- Bron, E. E., Smits, M., van der Flier, W. M., Vrenken, H., Barkhof, F., Scheltens, P., et al. (2015). Standardized evaluation of algorithms for computer-aided diagnosis of dementia based on structural MRI: the CADDementia challenge. *Neuroimage* 111, 562–579. doi: 10.1016/j.neuroimage.2015.01.048
- Cai, K. P., Xu, H., Guan, H., Zhu, W. L., Jiang, J. Y., Cui, Y., et al. (2017). Identification of early-stage Alzheimer's disease using sulcal morphology and other common neuroimaging indices. *PLoS ONE* 12:e0170875. doi: 10.1371/journal.pone.0170875
- Chawla, N. V., Bowyer, K. W., Hall, L. O., and Kegelmeyer, W. P. (2002). SMOTE: synthetic minority over-sampling technique. *J. Artif. Intell. Res.* 16, 321–357. doi: 10.1613/jair.953
- Cox, D. R. (1958). The regression analysis of binary sequences. *J. R. Stat. Soc.* 20, 215–242.
- Cui, Y., Sachdev, P. S., Lipnicki, D. M., Jin, J. S., Luo, S., Zhu, W., et al. (2012a). Predicting the development of mild cognitive impairment: a new use of pattern recognition. *Neuroimage* 60, 894–901. doi: 10.1016/j.neuroimage.2012.01.084
- Cui, Y., Wen, W., Lipnicki, D. M., Beg, M. F., Jin, J. S., Luo, S., et al. (2012b). Automated detection of amnesic mild cognitive impairment in community-dwelling elderly adults: a combined spatial atrophy and white matter alteration approach. *Neuroimage* 59, 1209–1217. doi: 10.1016/j.neuroimage.2011.08.013
- Cuingnet, R., Gerardin, E., Tessieras, J., Auzias, G., Lehericy, S., Habert, M. O., et al. (2011). Automatic classification of patients with Alzheimer's disease from structural MRI: a comparison of ten methods using the ADNI database. *Neuroimage* 56, 766–781. doi: 10.1016/j.neuroimage.2010.06.013
- Davatzikos, C., Bhatt, P., Shaw, L. M., Batmanghelich, K. N., and Trojanowski, J. Q. (2011). Prediction of MCI to AD conversion, via MRI, CSF biomarkers, and pattern classification. *Neurobiol. Aging* 32, 2322.e19–27. doi: 10.1016/j.neurobiolaging.2010.05.023
- Desikan, R. S., Segonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., et al. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage* 31, 968–980. doi: 10.1016/j.neuroimage.2006.01.021
- Dougherty, E. R., Sima, C., Hua, J. P., Hanczar, B., and Braga-Neto, U. M. (2010). Performance of error estimators for classification. *Curr. Bioinform.* 5, 53–67. doi: 10.2174/157489310790596385
- Dubey, R., Zhou, J. Y., Wang, Y. L., Thompson, P. M., Ye, J. P., and Alzheimer's Disease Neuroimaging Initiative. (2014). Analysis of sampling techniques for imbalanced data: an n=648 ADNI study. *Neuroimage* 87, 220–241. doi: 10.1016/j.neuroimage.2013.10.005
- Eskildsen, S. F., Coupe, P., Garcia-Lorenzo, D., Fonov, V., Pruessner, J. C., Collins, D. L., et al. (2013). Prediction of Alzheimer's disease in subjects with mild cognitive impairment from the ADNI cohort using patterns of cortical thinning. *Neuroimage* 65, 511–521. doi: 10.1016/j.neuroimage.2012.09.058
- Falahati, F., Westman, E., and Simmons, A. (2014). Multivariate Data Analysis and Machine Learning in Alzheimer's Disease with a Focus on Structural Magnetic Resonance Imaging. *J. Alzheimers Dis.* 41, 685–708. doi: 10.3233/JAD-131928
- Folstein, M., Folstein, S., and McHugh, P. (1975). "Mini-mental state": a practical method for grading the cognitive state of patients for the clinician. *J. Psychiatr. Res.* 12, 189–198. doi: 10.1016/0022-3956(75)90026-6
- Grundman, M., Petersen, R. C., Ferris, S. H., Thomas, R. G., Aisen, P. S., Bennett, D. A., et al. (2004). Mild cognitive impairment can be distinguished from Alzheimer disease and normal aging for clinical trials. *Arch. Neurol.* 61, 59–66. doi: 10.1001/archneur.61.1.59
- Guo, S., Lai, C., Wu, C., Cen, G., and Alzheimer's Disease Neuroimaging, I. (2017). Conversion discriminative analysis on mild cognitive impairment using multiple cortical features from MR Images. *Front. Aging Neurosci.* 9:146. doi: 10.3389/fnagi.2017.00146
- Guyon, I., Weston, J., Barnhill, S., and Vapnik, V. (2002). Gene selection for cancer classification using support vector machines. *Mach. Learn.* 46, 389–422. doi: 10.1023/A:1012487302797
- Hastie, T., Tibshirani, R., and Friedman, J. (2001). *The Elements of Statistical Learning*. New York, NY: Springer series in statistics.
- Hindmarch, I., Leheld, H., De, J. P., and Erzigkeit, H. (1998). The bayer activities of daily living scale (B-ADL). *Dement. Geriatr. Cogn. Disord.* 9 (Suppl. 2), 20–26. doi: 10.1159/000051195
- Huang, J., and Ling, C. X. (2005). Using, A. U. C., and accuracy in evaluating learning algorithms. *IEEE Trans. Knowl. Data Eng.* 17, 299–310. doi: 10.1109/TKDE.2005.50
- Huang, M., Yang, W., Feng, Q., Chen, W., and Alzheimer's Disease Neuroimaging Initiative. (2017). Longitudinal measurement and hierarchical classification framework for the prediction of Alzheimer's disease. *Sci. Rep.* 7:39880. doi: 10.1038/srep39880
- Jones, E., Oliphant, E., and Peterson, P. (2001). *SciPy: Open Source Scientific Tools for Python*. <http://www.scipy.org/> (Accessed Jan 11, 2017)
- Jung, N.-Y., Seo, S. W., Yoo, H., Yang, J.-J., Park, S., Kim, Y. J., et al. (2016). Classifying anatomical subtypes of subjective memory impairment. *Neurobiol. Aging* 48, 53–60. doi: 10.1016/j.neurobiolaging.2016.08.010
- Kochunov, P., Rogers, W., Mangin, J. F., and Lancaster, J. (2012). A Library of cortical morphology analysis tools to study development, aging and genetics of cerebral cortex. *Neuroinformatics* 10, 81–96. doi: 10.1007/s12021-011-9127-9
- Koepsell, T., and Monsell, S. (2011). Characterizing individuals who revert from mild cognitive impairment to normal or near-normal cognition. *Alzheimer's Dement.* 7:S539. doi: 10.1016/j.jalz.2011.05.1520
- Kriegeskorte, N., Simmons, W. K., Bellgowan, P. S., and Baker, C. I. (2009). Circular analysis in systems neuroscience: the dangers of double dipping. *Nat. Neurosci.* 12, 535–540. doi: 10.1038/nn.2303
- Lebedev, A. V., Westman, E., Van Westen, G. J., Kramberger, M. G., Lundervold, A., Aarsland, D., et al. (2014). Random forest ensembles for detection and prediction of Alzheimer's disease with a good between-cohort robustness. *Neuroimage Clin.* 6, 115–125. doi: 10.1016/j.nicl.2014.08.023
- Lebedeva, A. K., Westman, E., Borza, T., Beyer, M. K., Engedal, K., Aarsland, D., et al. (2017). MRI-based classification models in prediction of mild cognitive impairment and dementia in late-life depression. *Front. Aging Neurosci.* 9:13. doi: 10.3389/fnagi.2017.00013
- Li, Y., Wang, Y. P., Wu, G. R., Shi, F., Zhou, L. P., Lin, W. L., et al. (2012). Discriminant analysis of longitudinal cortical thickness changes in Alzheimer's disease using dynamic and network features. *Neurobiol. Aging* 33, 427.e15–30. doi: 10.1016/j.neurobiolaging.2010.11.008
- Liu, M., Zhang, J., Yap, P. T., and Shen, D. (2017). View-aligned hypergraph learning for Alzheimer's disease diagnosis with incomplete multi-modality data. *Med. Image Anal.* 36, 123–134. doi: 10.1016/j.media.2016.11.002
- Liu, S. Q., Liu, S. D., Cai, W. D., Che, H. Y., Pujol, S., Kikinis, R., et al. (2015). Multimodal Neuroimaging feature learning for multiclass diagnosis of Alzheimer's disease. *IEEE TBE* 62, 1132–1140. doi: 10.1109/TBME.2014.2372011
- Liu, T., Sachdev, P. S., Lipnicki, D. M., Jiang, J., Geng, G., Zhu, W., et al. (2013). Limited relationships between two-year changes in sulcal morphology and

- other common neuroimaging indices in the elderly. *Neuroimage* 83, 12–17. doi: 10.1016/j.neuroimage.2013.06.058
- Lopez, O. L., Becker, J. T., Chang, Y. F., Sweet, R. A., DeKosky, S. T., Gach, M. H., et al. (2012). Incidence of mild cognitive impairment in the pittsburgh cardiovascular health study-cognition study. *Neurology* 79, 1599–1606. doi: 10.1212/WNL.0b013e31826e25f0
- Maclin, R., and Opitz, D. (1999). Popular ensemble methods: an empirical study. *J. Artif. Intell. Res.* 11, 169–198.
- Macqueen, J. (1967). “Some methods for classification and analysis of multivariate observations,” in *Proceedings of Berkeley Symposium on Mathematical Statistics and Probability* (Berkeley, CA: University of California Press), 281–297.
- Mayo, C. D., Mazerolle, E. L., Ritchie, L., Fisk, J. D., Gawryluk, J. R., and Alzheimer's Disease Neuroimaging Initiative. (2017). Longitudinal changes in microstructural white matter metrics in Alzheimer's disease. *Neuroimage Clin.* 13, 330–338. doi: 10.1016/j.nicl.2016.12.012
- McEvoy, L. K., Holland, D., Hagler, D. J. Jr., Fennema-Notestine, C., Brewer, J. B., Dale, A. M., et al. (2011). Mild cognitive impairment: baseline and longitudinal structural MR imaging measures improve predictive prognosis. *Radiology* 259, 834–843. doi: 10.1148/radiol.11101975
- Min, R., Wu, G., Cheng, J., Wang, Q., Shen, D., and Alzheimer's Disease Neuroimaging Initiative. (2014). Multi-atlas based representations for Alzheimer's disease diagnosis. *Hum. Brain Mapp.* 35, 5052–5070. doi: 10.1002/hbm.22531
- Misra, C., Fan, Y., and Davatzikos, C. (2009). Baseline and longitudinal patterns of brain atrophy in MCI patients, and their use in prediction of short-term conversion to AD: results from ADNI. *Neuroimage* 44, 1415–1422. doi: 10.1016/j.neuroimage.2008.10.031
- Moradi, E., Pepe, A., Gaser, C., Huttunen, H., Tohka, J., and Alzheimer's Disease Neuroimaging, I. (2015). Machine learning framework for early MRI-based Alzheimer's conversion prediction in MCI subjects. *Neuroimage* 104, 398–412. doi: 10.1016/j.neuroimage.2014.10.002
- Mwangi, B., Tian, T. S., and Soares, J. C. (2014). A review of feature reduction techniques in neuroimaging. *Neuroinformatics* 12, 229–244. doi: 10.1007/s12021-013-9204-3
- Pandya, S. Y., Clem, M. A., Silva, L. M., and Woon, F. L. (2016). Does mild cognitive impairment always lead to dementia? A review. *J. Neurol. Sci.* 369, 57–62. doi: 10.1016/j.jns.2016.07.055
- Patenaude, B., Smith, S. M., Kennedy, D. N., and Jenkinson, M. (2011). A Bayesian model of shape and appearance for subcortical brain segmentation. *Neuroimage* 56, 907–922. doi: 10.1016/j.neuroimage.2011.02.046
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al. (2011). Scikit-learn: machine learning in python. *J. Mach. Learn. Res.* 12, 2825–2830.
- Penttilä, J., Paillere-Martinot, M. L., Martinot, J. L., Ringuenet, D., Wessa, M., Houenou, J., et al. (2009). Cortical folding in patients with bipolar disorder or unipolar depression. *J. Psychiatry Neurosci.* 34, 127–135.
- Pereira, F., Mitchell, T., and Botvinick, M. (2009). Machine learning classifiers and fMRI: a tutorial overview. *Neuroimage* 45, S199–S209. doi: 10.1016/j.neuroimage.2008.11.007
- Petersen, R. C. (2004). Mild cognitive impairment as a diagnostic entity. *J. Intern. Med.* 256, 183–194. doi: 10.1111/j.1365-2796.2004.01388.x
- Petersen, R. C. (2010). Does the source of subjects matter? Absolutely! *Neurology* 74, 1754–1755. doi: 10.1212/WNL.0b013e3181e533e7
- Reuter, M., and Fischl, B. (2011). Avoiding asymmetry-induced bias in longitudinal image processing. *Neuroimage* 57, 19–21. doi: 10.1016/j.neuroimage.2011.02.076
- Reuter, M., Schmansky, N. J., Rosas, H. D., and Fischl, B. (2012). Within-subject template estimation for unbiased longitudinal image analysis. *Neuroimage* 61, 1402–1418. doi: 10.1016/j.neuroimage.2012.02.084
- Risacher, S. L., Shen, L., West, J. D., Kim, S., McDonald, B. C., Beckett, L. A., et al. (2010). Longitudinal MRI atrophy biomarkers: relationship to conversion in the ADNI cohort. *Neurobiol. Aging* 31, 1401–1418. doi: 10.1016/j.neurobiolaging.2010.04.029
- Rivière, D., Geffroy, D., Denghien, I., Souedet, N., and Cointepas, Y. (2009). BrainVISA: an extensible software environment for sharing multimodal neuroimaging data and processing tools. *Neuroimage* 47, S163–S163. doi: 10.1016/S1053-8119(09)71720-3
- Rivière, D., Mangin, J. F., Papadopoulos-Orfanos, D., Martinez, J. M., Frouin, V., and Regis, J. (2002). Automatic recognition of cortical sulci of the human brain using a congregation of neural networks. *Med. Image Anal.* 6, 77–92. doi: 10.1016/S1361-8415(02)00052-X
- Roberts, R. O., Knopman, D. S., Mielke, M. M., Cha, R. H., Pankratz, V. S., Christianson, T. J. H., et al. (2014). Higher risk of progression to dementia in mild cognitive impairment cases who revert to normal. *Neurology* 82, 317–325. doi: 10.1212/WNL.0000000000000055
- Rosen, W. G., Mohs, R. C., and Davis, K. L. (1984). A new rating scale for Alzheimer's disease. *Am. J. Psychiatry* 141, 1356–1364. doi: 10.1176/ajp.141.11.1356
- Rusinek, H., De Santi, S., Frid, D., Tsui, W. H., Tarshish, C. Y., Convit, A., et al. (2003). Regional brain atrophy rate predicts future cognitive decline: 6-year longitudinal MR imaging study of normal aging. *Radiology* 229, 691–696. doi: 10.1148/radiol.2293021299
- Sabuncu, M. R., Desikan, R. S., Sepulcre, J., Yeo, B. T., Liu, H., Schmansky, N. J., et al. (2011). The dynamics of cortical and hippocampal atrophy in Alzheimer disease. *Arch. Neurol.* 68, 1040–1048. doi: 10.1001/archneurol.2011.167
- Sachdev, P. S., Brodaty, H., Reppermund, S., Kochan, N. A., Trollor, J. N., Draper, B., et al. (2010). The Sydney Memory and Ageing Study (MAS): methodology and baseline medical and neuropsychiatric characteristics of an elderly epidemiological non-demented cohort of Australians aged 70–90 years. *Int. Psychogeriatr.* 22, 1248–1264. doi: 10.1017/S1041610210001067
- Schoonheim, M. M., Hulst, H. E., Brandt, R. B., Strik, M., Wink, A. M., Uitendag, B. M. J., et al. (2015). Thalamus structure and function determine severity of cognitive impairment in multiple sclerosis. *Neurology* 84, 776–783. doi: 10.1212/WNL.0000000000001285
- Shao, J., Myers, N., Yang, Q., Feng, J., Plant, C., Bohm, C., et al. (2012). Prediction of Alzheimer's disease using individual structural connectivity networks. *Neurobiol. Aging* 33, 2756–2765. doi: 10.1016/j.neurobiolaging.2012.01.017
- Shi, F., Liu, B., Zhou, Y., Yu, C., and Jiang, T. (2010). Hippocampal volume and asymmetry in mild cognitive impairment and Alzheimer's disease: meta-analyses of MRI studies. *Hippocampus* 19, 1055–1064. doi: 10.1002/hipo.20573
- Sun, Z. Y., Riviere, D., Poupon, F., Regis, J., and Mangin, J. F. (2007). Automatic inference of sulcus patterns using 3D moment invariants. *Med. Image Comput. Comput. Assist. Interv.* 4791, 515–522. doi: 10.1007/978-3-540-75757-3_63
- Tabert, M. H., Manly, J. J., Liu, X., Pelton, G. H., Rosenblum, S., Jacobs, M., et al. (2006). Neuropsychological prediction of conversion to Alzheimer disease in patients with mild cognitive impairment. *Arch. Gen. Psychiatry* 63, 916–924. doi: 10.1001/archpsyc.63.8.916
- Tohka, J., Moradi, E., Huttunen, H., and Alzheimer's Disease Neuroimaging Initiative (2016). Comparison of feature selection techniques in machine learning for anatomical brain MRI in dementia. *Neuroinformatics* 14, 279–296. doi: 10.1007/s12021-015-9292-3
- Trojanowski, J. Q., Vandevertichele, H., Korecka, M., Clark, C. M., Aisen, P. S., Petersen, R. C., et al. (2010). Update on the biomarker core of the Alzheimer's disease neuroimaging initiative subjects. *Alzheimer's Dement.* 6, 230–238. doi: 10.1016/j.jalz.2010.03.008
- Trzepacz, P. T., Yu, P., Sun, J., Schuh, K., Case, M., Witte, M. M., et al. (2014). Comparison of neuroimaging modalities for the prediction of conversion from mild cognitive impairment to Alzheimer's dementia. *Neurobiol. Aging* 35, 143–151. doi: 10.1016/j.neurobiolaging.2013.06.018
- Vapnik, V. N. (1995). *The Nature of Statistical Learning Theory*. New York, NY: Springer.
- Walt, S. V. D., Colbert, S. C., and Varoquaux, G. (2011). The NumPy array: a structure for efficient numerical computation. *Comput. Sci. Eng.* 13, 22–30. doi: 10.1109/MCSE.2011.37
- Wee, C.-Y., Yap, P.-T., Li, W., Denny, K., Brownhyke, J. N., Potter, G. G., et al. (2011). Enriched white matter connectivity networks for accurate identification of MCI patients. *Neuroimage* 54, 1812–1822. doi: 10.1016/j.neuroimage.2010.10.026
- Wee, C.-Y., Yap, P.-T., Zhang, D., Denny, K., Brownhyke, J. N., Potter, G. G., et al. (2012). Identification of MCI individuals using structural and functional connectivity networks. *Neuroimage* 59, 2045–2056. doi: 10.1016/j.neuroimage.2011.10.015
- Wen, W., Sachdev, P. S., Li, J. J., Chen, X. H., and Anstey, K. J. (2009). White Matter Hyperintensities in the forties: their prevalence and topography in

- an epidemiological sample aged 44–48. *Hum. Brain Mapp.* 30, 1155–1167. doi: 10.1002/hbm.20586
- Whitwell, J. L., Petersen, R. C., Negash, S., Weigand, S. D., Kantarci, K., Ivnik, R. J., et al. (2007). Patterns of atrophy differ among specific subtypes of mild cognitive impairment. *Arch. Neurol.* 64, 1130–1138. doi: 10.1001/archneur.64.8.1130
- Winblad, B., Palmer, K., Kivipelto, M., Jelic, V., Fratiglioni, L., Wahlund, L. O., et al. (2004). Mild cognitive impairment - beyond controversies, towards a consensus: report of the International Working Group on Mild Cognitive Impairment. *J. Intern. Med.* 256, 240–246. doi: 10.1111/j.1365-2796.2004.01380.x
- Yun, H. J., Kwak, K., and Lee, J. M. (2015). Multimodal discrimination of Alzheimer's Disease based on regional cortical atrophy and hypometabolism. *PLoS ONE* 10:e0129250. doi: 10.1371/journal.pone.0129250
- Zhang, D., Shen, D., and Alzheimer's Disease Neuroimaging Initiative (2012a). Predicting future clinical changes of MCI patients using longitudinal and multimodal biomarkers. *PLoS ONE* 7:e33182. doi: 10.1371/journal.pone.0033182
- Zhang, D., Wang, Y., Zhou, L., Yuan, H., and Shen, D. (2011). Multimodal classification of Alzheimer's disease and mild cognitive impairment. *Neuroimage* 55, 856–867. doi: 10.1016/j.neuroimage.2011.01.008
- Zhang, H. B., Sachdev, P. S., Wen, W., Kochan, N. A., Crawford, J. D., Brodaty, H., et al. (2012b). Gray matter atrophy patterns of mild cognitive impairment subtypes. *J. Neurol. Sci.* 315, 26–32. doi: 10.1016/j.jns.2011.12.011
- Zhang, J., Liu, M., An, L., Gao, Y., and Shen, D. (2017). Alzheimer's disease diagnosis using landmark-based features from longitudinal structural mr images. *IEEE J. Biomed. Health Informat.* doi: 10.1109/JBHI.2017.2704614. [Epub ahead of print].
- Zhou, J., Liu, J., Narayan, V. A., Ye, J., and Alzheimer's Disease Neuroimaging Initiative (2013). Modeling disease progression via multi-task learning. *Neuroimage* 78, 233–248. doi: 10.1016/j.neuroimage.2013.03.073

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.

Copyright © 2017 Guan, Liu, Jiang, Tao, Zhang, Niu, Zhu, Wang, Cheng, Kochan, Brodaty, Sachdev and Wen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Combining SPECT and Quantitative EEG Analysis for the Automated Differential Diagnosis of Disorders with Amnestic Symptoms

Yvonne Höller^{1*}, Arne C. Bathke², Andreas Uhl³, Nicolas Strobl¹, Adelheid Lang⁴, Jürgen Bergmann¹, Raffaele Nardone^{1,5,6}, Fabio Rossini¹, Harald Zauner⁷, Margarita Kirschner¹, Amirhossein Jahanbekam⁸, Eugen Trinkä^{1,5} and Wolfgang Staffen¹

¹ Department of Neurology, Christian Doppler Medical Centre and Centre for Cognitive Neuroscience, Paracelsus Medical University of Salzburg, Salzburg, Austria, ² Department of Mathematics, Paris Lodron University of Salzburg, Salzburg, Austria, ³ Multimedia Signal Processing and Security Lab, Department of Computer Sciences, Paris Lodron University of Salzburg, Salzburg, Austria, ⁴ Department of Psychology, Centre for Cognitive Neuroscience, Paris Lodron University of Salzburg, Salzburg, Austria, ⁵ Spinal Cord Injury and Tissue Regeneration Center, Paracelsus Medical University of Salzburg, Salzburg, Austria, ⁶ Department of Neurology, Franz Tappeiner Hospital, Merano, Italy, ⁷ Cardiovascular and Neurological Rehabilitation Center, Großgmain, Austria, ⁸ Department of Epileptology, University of Bonn, Bonn, Germany

OPEN ACCESS

Edited by:

Stefan Teipel,
German Center for Neurodegenerative
Diseases (HZ), Germany

Reviewed by:

Martin Dyrba,
German Center for Neurodegenerative
Diseases (HZ), Germany
Christian Salvatore,
Institute of Molecular Bioimaging and
Physiology (CNR), Italy
Alessia Sarica,
Institute of Molecular Bioimaging and
Physiology (CNR), Italy
Lora Minkova,
Universitätsklinikum Freiburg,
Germany

*Correspondence:

Yvonne Höller
y.hoeller@salk.at

Received: 31 March 2017

Accepted: 23 August 2017

Published: 07 September 2017

Citation:

Höller Y, Bathke AC, Uhl A, Strobl N, Lang A, Bergmann J, Nardone R, Rossini F, Zauner H, Kirschner M, Jahanbekam A, Trinkä E and Staffen W (2017) Combining SPECT and Quantitative EEG Analysis for the Automated Differential Diagnosis of Disorders with Amnestic Symptoms. *Front. Aging Neurosci.* 9:290. doi: 10.3389/fnagi.2017.00290

Single photon emission computed tomography (SPECT) and Electroencephalography (EEG) have become established tools in routine diagnostics of dementia. We aimed to increase the diagnostic power by combining quantitative markers from SPECT and EEG for differential diagnosis of disorders with amnestic symptoms. We hypothesize that the combination of SPECT with measures of interaction (*connectivity*) in the EEG yields higher diagnostic accuracy than the single modalities. We examined 39 patients with Alzheimer's dementia (AD), 69 patients with depressive cognitive impairment (DCI), 71 patients with amnestic mild cognitive impairment (aMCI), and 41 patients with amnestic subjective cognitive complaints (aSCC). We calculated 14 measures of interaction from a standard clinical EEG-recording and derived graph-theoretic network measures. From regional brain perfusion measured by 99mTc-hexamethyl-propylene-aminoxime (HMPAO)-SPECT in 46 regions, we calculated relative cerebral perfusion in these patients. Patient groups were classified pairwise with a linear support vector machine. Classification was conducted separately for each biomarker, and then again for each EEG- biomarker combined with SPECT. Combination of SPECT with EEG-biomarkers outperformed single use of SPECT or EEG when classifying aSCC vs. AD (90%), aMCI vs. AD (70%), and AD vs. DCI (100%), while a selection of EEG measures performed best when classifying aSCC vs. aMCI (82%) and aMCI vs. DCI (90%). Only the contrast between aSCC and DCI did not result in above-chance classification accuracy (60%). In general, accuracies were higher when measures of interaction (i.e., connectivity measures) were applied directly than when graph-theoretical measures were derived. We suggest that quantitative analysis of EEG and machine-learning techniques can support differentiating AD, aMCI, aSCC, and DCC, especially when being combined with imaging methods such as SPECT. Quantitative analysis of EEG connectivity could become an integral part for early differential diagnosis of cognitive impairment.

Keywords: SPECT, EEG connectivity, dementia, depression with cognitive impairment, mild cognitive impairment, subjective cognitive complaints

1. INTRODUCTION

Mild cognitive impairment (MCI) is common in the elderly population and can be stable or convert to Alzheimer's disease (AD) (Winblad et al., 2004; Gauthier et al., 2006). Estimated 47.5 million people suffer from dementia worldwide, and it is estimated that this number will triple by 2050 (World Health Organization, 2016). The WHO reports an estimate of US \$604 billion of total global costs associated with dementia. Early differential diagnosis of MCI, subjective cognitive complaints (SCC), and depressive cognitive impairment (DCI) would pave the way for new therapeutic programs, possibly reducing the overall burden of memory disorders and improving quality of life of these patients (DeKosky and Marek, 2003). Because of the various aetiologies and pathologic processes that may lead to memory impairments it is suggested that a combination of several biomarkers is necessary to provide an early diagnosis of AD in the various phases and variations of the disease (Scheltens et al., 1997; DeKosky and Marek, 2003; Wurtman, 2015).

The National Institute of Neurological and Communicative Diseases and Stroke/Alzheimer's Disease and Related Disorders Association (NINCDS-ADRDA) has proposed clinical criteria for the diagnosis of probable AD (McKhann et al., 1984). For an early detection it is not enough to use neuropsychological tests alone since SCC are—by definition—not detectable by these diagnostic procedures, i.e., they are experienced subjectively, only. A patient may suffer from impairment and notice the change. However, a neuropsychological test indicates only whether the patient scores lower than the reference group that was used to standardize the test. When a patient performs well above average throughout his life and experiences a loss because of beginning MCI, he might still perform within the normal range, despite having subjectively noticed the objective decline. In turn, the diagnosis of MCI is still a challenge for neuropsychologists (Ladeira et al., 2009; Lopez, 2013; Rentz et al., 2013). In addition, some of the physiological features that differentiate several types of dementia cannot be assessed with behavioral tests. In the following, we want to outline two diagnostic modalities that might complement each other and thus, are hypothesized to contribute to the differential diagnosis of disorders with amnesic symptoms.

Single Photon Emission Computer Tomography (SPECT) is complementary to clinical assessment (Farid et al., 2011). The measured activity, i.e., the perfusion, can be quantified by volumetric analysis of activated brain regions either manually, semi-automatically, or fully automatically, such as with statistical parametric mapping (SPM) (Friston, 1995; Van Heertum et al., 2009), specifically for differentiating AD from different types of dementia (Kemp et al., 2005). By providing functional information, early stages of cognitive impairment can be identified and differentiation between MCI, AD, and/or other types of cognitive dysfunction can be achieved (Bonte et al., 1990; Talbot et al., 1998; Staffen et al., 2006, 2009; Van Heertum et al., 2009; Farid et al., 2011). Specifically, 99mTc-hexamethylpropylene- aminoxime (HMPAO)-SPECT seems to be sensitive to cognitive impairment, AD and prodromal stages of AD (e.g., Goldenberg et al., 1989; Frisoni et al., 2014; Swan et al., 2015; Valotassiou et al., 2015). Even when contrasting patients

with subjective memory complaints to patients with memory impairment, HMPAO SPECT can be sensitive to cerebral hypoperfusion (Banzo et al., 2011). However, not all studies fully support the usefulness of SPECT for differential diagnosis of disorders with amnesic symptoms (Barnes et al., 2000; Kaneko et al., 2004). Therefore, we suggest combination of SPECT with another physiological modality.

Characteristics from the electroencephalogram (EEG) distinguish patients with AD from MCI and patients with MCI from healthy subjects (see Rossini et al., 2007; Dauwels et al., 2010, for a review). The classical clinical finding is the slow alpha rhythm, which can be quantified as an increase of slow activity; Fast Fourier transform shows a relative increase of activity below 8 Hz and a decrease above this range. The use of the EEG in the assessment of AD dates back to 1952 (see Brenner, 1999, for review). Today it is assumed that the shift toward lower frequencies is possibly caused by perturbations in synchronization and decreased neural complexity (Cantero et al., 2009). Synchronization may be increased or decreased in MCI depending on frequency range, type of analysis, and regions being assessed (Jelic et al., 2000; Koenig et al., 2005; Stam, 2005; Babiloni et al., 2006). Interactions between neural signals are at the forefront of current neuroscientific research, which is also emphasized by the most recent name for this phenomenon: connectomics (Sporns, 2015). The assessment of the connectome has attracted particularly great interest with regard to brain disorders (Fornito et al., 2015). In MCI, interaction between EEG-signals (today, mostly known as connectivity, Aertsen and Preissl, 1991) was found to be a reliable marker for cerebral reserve capacity (Teipel et al., 2016), response to interventions (Klados et al., 2016), and to monitor disease progression (see for recent examples Dimitriadis et al., 2015; Hatz et al., 2015; Wurtman, 2015; Babiloni et al., 2016; Miraglia et al., 2016; Vecchio et al., 2016). Among the plethora of measures indicating interactions between brain regions it is neither clear which ones are preferable over others for diagnostic purposes, nor do we know whether the integration of these measures in to graph-theoretic network characteristics could be a viable method for feature reduction. Therefore, it is recommendable to compare different approaches for characterization of EEG interactions (Lehnertz, 2011). However, because of the low spatial resolution of the EEG, we suggest that it should be combined with neuroimaging in order to yield a full picture of altered brain activity in amnesic disorders.

While it was suggested that the combination of different modalities would contribute to the diagnostic process (Scheltens et al., 1997; DeKosky and Marek, 2003; Wurtman, 2015), only little research was done on the combination of SPECT with EEG. Some studies tried to associate EEG and cerebral perfusion values in patients with AD (Gueguen et al., 1991; Frölich et al., 1992; Sloan et al., 1995). EEG slowing is associated with reduced blood flow in temporo-parietal regions of AD patients (Kwa et al., 1993; Sloan et al., 1994). Degrees of interhemispheric asymmetry of EEG and SPECT are concordant in patients with AD (Montplaisir et al., 1996). Global decrease in cerebral blood flow correlates with a posterior shift of the topographical alpha-centroids (Müller et al., 1997). Power in the

EEG delta and alpha frequency ranges correlates with perfusion level in parietal regions and power in the EEG delta range with hippocampal perfusion level of AD patients (Rodriguez et al., 1999). In addition to these correlative studies, some evidence points to a possible complementary use of SPECT and EEG. There is an interaction between alterations in event related potentials recorded with EEG and changes of cerebral blood flow characterized by HMPAO SPECT in AD (Gungor et al., 2005). Specifically, EEG changes take place at earlier stages of the condition than the changes in cerebral blood flow.

No study so far examined the additional value of merging information from advanced EEG measures of interaction and cerebral blood flow measured by SPECT in order to differentiate patients with different types of amnesic syndromes at different stages of AD. We hypothesize that the combined analysis of cerebral perfusion as indicated by HMPAO SPECT and quantitative measures of interaction from the EEG by applying modern methods of data analysis will increase the diagnostic accuracy.

In this study, we assessed the significance of combining EEG measures of interaction or graph-theoretical network characteristics with SPECT perfusion values for differential diagnosis of amnesic SCC (aSCC), amnesic MCI (aMCI), AD, or DCI. Specific expectations about characteristics from the EEG or SPECT that could be most distinctive are restricted to the slowing of the EEG networks, which is more prominent at more advanced stages of cognitive decline, as well as parietal and hippocampal hypoperfusion. Therefore, we applied a machine-learning approach that should identify the most distinctive combination of features from both modalities to pairwise group classifications.

2. MATERIALS AND METHODS

2.1. Ethics

The study was conducted as a retrospective data analysis. Several years after the examination of the patients had been performed, we analyzed routinely recorded EEG, SPECT, and clinical data. The local Ethics Committee (Ethics Commission Salzburg/Ethikkommission Land Salzburg) confirmed that there are no ethical concerns with respect to this study.

2.2. Subjects

We selected 220 consecutive patients from the data repositories at the Department of Neurology, Paracelsus Medical University Salzburg, Austria, which were examined in the memory clinic between June 2007 and March 2011. Diagnosis of aSCC, aMCI, AD, or DCI was assigned at the time of examinations, based on multimodal assessment in the memory clinic of our hospital, including a neurological and neuropsychological examination [German version of the hospital anxiety and depression scale; HADS-D (Zigmond and Snaith, 1983; Herrmann-Lingen et al., 2007), test battery of the Consortium to Establish a Registry for Alzheimer's Disease; CERAD (Morris et al., 1989; Welsh et al., 1994; Thalman et al., 2000), including a slightly modified version of the mini-mental state examination MMSE

by Folstein (Folstein et al., 1975), and in addition (known as the CERAD-Plus tests), the trail making test (Reitan, 1979), and the test for phonematic verbal fluency (Spreeen and Benton, 1977)]. The examination included routine laboratory investigations, supplemented by determination of thyroid parameters, internal diagnostics (including electrocardiogram, ECG), cranial computed tomography (CCT), ultrasonographic examination of the carotid and vertebrobasilar arteries, and a cerebral perfusion SPECT scan. The latter was exclusively employed in the differential diagnosis of AD vs. Lewy body dementia, frontotemporal dementia, and vascular dementia based on visually evaluated different patterns of perfusion disturbance. An EEG was recorded in order to disclose epileptic activity.

The diagnosis was assigned by the medical doctor according to the results of the described multimodal examination according to the criteria of Petersen (Petersen et al., 1999). Specifically, we conformed to the definition of aMCI and aSCC where amnesic aMCI equals to level three and patients with aSCC equals to level two of the global deterioration scale for aging and dementia (Winblad et al., 2004; Gauthier et al., 2006). Most importantly, the diagnosis of aMCI and aSCC indicates that the complaints and/or deficits were detectable only in the memory domain, and not on other cognitive subscales.

Patients with DCI were treated with antidepressants after the examinations clarified the diagnosis. However, not all of them were drug-naïve at the time of examination since antidepressants are commonly prescribed in the elderly by the general practitioner in order to treat self-reported mood complaints and sleep disorders.

Please note that the diagnosis did not include quantitative assessment of SPECT and EEG as done for the present work. Thus, the original diagnosis of memory impairment was not based on the quantitative analysis as described in the subsequent sections.

2.3. SPECT Examination

The SPECT examination was performed under quiet conditions (relaxed lying in quiet surroundings and dimmed light 10 min before the injection and during the whole time of the examination), with ^{99m}Tc-hexamethyl-propylene- aminoxime (HMPAO, Ceretec, Amersham, UK) serving as perfusion tracer at a dose of 740 MBq. Perfusion was measured 20 min after injection with a three-headed gamma camera (Prism 3000, Picker International, Imaging Division, Cleveland, OH) over 35133815min (3° for 40 steps, i.e., in sum 120°). Datasets were corrected for scatter and attenuation, reconstructed using filtered back projection and displayed as a set of 20 slices using a 128 × 128 matrix. Attenuation correction was applied at the time of reconstruction using Chang's first-order approximation of linear attenuation ($\mu = 0.09/\text{cm}$), within an elliptical contour fitted to every slice of the brain (Chang, 1978).

2.4. SPECT Analysis

For analysis of SPECT data a region of interest (ROI) regionalization was performed automatically to assess relative blood flow (cerebellar ratios) of 46 brain regions. Data

were quantified semiautomatically, using the HERMES BRASS Software package (Hermes Medical Solutions, Stockholm, Sweden) which spatially co-registered the image data to an anatomically standardized, stereotactic template consisting of scans of 35 healthy volunteers. Data were count-normalized by the cerebellar count rate and compared to the normal population voxel-by-voxel, as well as on a regional basis. The region map used therefore was predefined using a normal T1-weighted MRI scan co-registered to the normal template.

The regions for which we obtained relative blood flow were cerebellar cortex, cerebellar white matter, nucleus lentiformis, nucleus caudatus, thalamus, sensorimotor cortex, occipital cortex, superior parietal lobule, anterior dorsal frontal region, posterior dorsal frontal region, anterior orbital frontal region, posterior orbital cortex, parietotemporal cortex, medial temporal lobe, lateral temporal lobe, posterior temporal lobe, temporal pole, insular cortex, anterior cingulate gyrus, posterior cingulate gyrus, anterior subcortical region, posterior subcortical region, each of these separately for left and right hemisphere, and in addition one region including pons and midbrain and one region including other subcortical regions. Thus, in sum, the SPECT-feature vector had a length of 46 values.

2.5. EEG Data Registration

EEG was recorded in a quiet room with a clinical standard electrode montage (10–20 Stellate Harmonie Routine EEG System by Natus, 21 channels placed in standard 10–20 EEG system) ground on Fpz, reference on Fcz, with additional earlobe-electrodes for re-referencing, and a sampling rate of 200 Hz. Impedances were kept below 10 k Ω . The EEG recording started with artifact provocation/calibration procedures. Subsequently standard intermittent light stimulation and hyperventilation were performed. Afterwards, the patients were asked to relax with eyes closed.

2.6. EEG data extraction

From a period of wakefulness with eyes closed a trained neuroscientist (co-author AL) extracted 3 min of EEG that were free of artifacts, e.g., muscle, eye, movement, etc. Data analysis was conducted for 17 electrodes: F3, F4, C3, C4, P3, P4, O1, O2, F7, F8, T3, T4, T5, T6, Fz, Cz, and Pz. The pre-selected EEG segments were exported into EDF and imported for further processing to Matlab[®] (release R2016b, The Mathworks, Massachusetts, USA).

2.7. Feature Extraction

We estimated a set of measures of interaction between all of the 17 selected electrodes (i.e., channels). The estimation was performed for each of the participants. The measures were calculated with the functions `mvfreqz.m` and `mvar.m` from the BioSig toolbox (Schlögl and Brunner, 2008) with model order 100 (i.e., equaling half of the sampling rate allowing to model at least one full oscillation beginning from 2 Hz). To estimate the multivariate autoregressive model we used partial correlation estimation with unbiased covariance estimates (Marple, 1987), which was found to be the most accurate estimation method according to Schlögl (2006). The

model was then transformed from the time-domain into the z -domain and the f -domain, which yielded accordingly two transfer functions. The multivariate parameters in the frequency domain that could be derived from these transfer functions were computed for 1 Hz frequency steps between 2 and 80 Hz. The following measures were extracted: auto- and cross-spectrum (S), direct causality (DC), transfer function (h), transfer function polynomial (Af), real valued coherence (COH), complex coherence (iCOH), partial coherence (pCOH), partial directed coherence (PDC), partial directed coherence factor (PDCF), generalized partial directed coherence (GPDC), directed transfer function (DTF), direct directed transfer function (dDTF), full frequency directed transfer function (ffDTF), and Geweke's Granger causality (GGC). A description of these measures and the references can be retrieved in the Supplementary Material Section.

Before statistically determining and evaluating the network characteristics, we averaged the network characteristics in classical frequency ranges delta (2–4 Hz), theta (5–7 Hz), alpha (8–13 Hz), beta (14–30 Hz), and gamma (31–80 Hz).

Finally, we derived graph theoretical measures for each of the listed measures by use of the Brain Connectivity Toolbox (Rubinov and Sporns, 2010). Thus, we calculated global network parameters from the connection matrices in each frequency range obtained from each of the multivariate parameters: assortativity, efficiency, clustering coefficient, modularity, and transitivity. For more details and references of these values we refer to the Supplementary Material.

2.8. Feature Vectors

Classification and cross-validation was repeated for the following scenarios, which can be described by their respective feature vectors including the following:

1. **EEG single:** each EEG measure individually, that is, feature vector optimization and classification was repeated for each EEG measure, where in every case the initial feature vector was formed as a concatenation of all measures of interaction for each electrode combination and all frequencies
2. **SPECT:** SPECT perfusion values of the 46 brain regions formed the initial feature vector
3. **EEG single + SPECT:** a combination of each EEG measure from (1) with the SPECT perfusion values from (2), separately for each EEG measure
4. **EEG merged:** a combination of all optimized feature vectors of the EEG measures from (1) in one feature vector formed the initial feature vector
5. **EEG merged + SPECT:** a combination of all optimized feature vectors of the EEG measures from (1) in one feature vector as in (4), combined with the optimized feature vector of SPECT perfusion values from (2) formed the initial feature vector
6. **EEG graph:** a combination of all EEG measures from (1) converted to graph-theoretic measures
7. **EEG graph + SPECT:** a combination of all EEG measures from (1) converted to graph-theoretic measures as in (6) combined with the SPECT perfusion values from (2).

2.9. Classification analysis

We performed pair-wise classification of all four groups, resulting in 6 group comparisons.

Supervised learning for classification typically includes a training and a testing step, with disjunctive samples for these two steps. That is, the data is divided into two subsets, one is used only for training, and one only for testing according to a defined strategy of cross-validation. The algorithm learns with the training data according to the properties of the samples and their labels, that is, the diagnosis. The result of this learning step is a model that allows to distinguish the members of the groups. In the second step, the algorithm is given only the data of the testing subset, but without the labels. The task is now to predict the correct labels based on the model that was built in the learning step and the data. In order to assess the quality of the classification, the correctness of the predicted labels can be evaluated.

We decided to use support vector machines for classification, because they deal with non-linear properties of the data even when a linear kernel is used. When data are only non-linearly separable, the data is mapped into a feature space in which the linear separating hyperplane can be used. We performed a classification in the sense of supervised learning with a linear kernel function (dot product) and quadratic programming in order to find the separating hyperplane, resulting in a 2-norm soft-margin support vector machine, by using the MATLAB functions `svmtrain` and `svmclassify` from the statistics and machine learning toolbox.

2.10. Feature Subset Selection

We performed a nested cross-validation with 3 layers with feature vector optimization, that is, feature subset selection, for each group comparison as illustrated in **Figure 1**.

Because of the high dimensionality of the data, we implemented a feature subset selection procedure. This procedure is used for two purposes. First, it is known that when the length of the feature vector exceeds the size of the sample, it can cause artificially high accuracies due to overfitting. Thus, shortening the feature vector to a length that is smaller than the training sample prevents us from running into the small sample size problem. This is easily the case for the EEG feature vectors, because then the length of the feature vector is up to $17 \times 17 \times 5$. Second, a long feature vector with uninformative features prevents the machine learning algorithm from finding a good solution. Therefore, the shortest possible feature vector should be found in the sense of a feature vector optimization. Because the maximally available features for SPECT was 46, we limited the maximally acceptable length of the feature vector to $0.9 \cdot 46 \sim 41$ entries. This is well below the smallest sample when combining the two smallest groups of AD ($N = 39$) and aSCC ($N = 41$), where the training sample in the outermost cross validation was $0.9 \cdot 80 = 72$.

As described in **Figure 1**, the classification and feature subset selection procedure was done in a nested design with 3 layers. We implemented an outer layer as a division of the data into 10% of the data for testing the resulting model, and 90% for feature vector optimization and cross validation, i.e., submitted

to the middle layer. The middle layer is a first inner loop, implemented with 10-fold cross-validation. This loop aims to estimate the consistency of selected features, since each run yields a different feature vector. The inner layer is a second, thus, nested inner loop, again with 10-fold cross-validation in order to perform adequate feature subset selection. So-called k -fold cross-validation consist of k repetitions of leaving out N/k samples as the training set, while the remaining $N - (N/k)$ samples are used during the training step.

All subsets were drawn in order to maintain the original proportion of the two groups.

Thus, the whole algorithm can be described as follows:

1. First, 10% of the sample were excluded as the outer-layer test set for the final validation step in the outer layer, while the remaining 90% of patients were used as the outer-layer training set, submitted to the next step
2. The outer-layer training set obtained from the outer loop was divided into 10 equal sized subsets, each one maintaining the proportion of group sizes from the original sample. For each of these 10 sets, the following steps were repeated in order to yield a greedy stepwise feature selection with forward search:
 - a. This set was left out, the other 9 sets were merged to form the middle-layer training set.
 - b. A t -test for the middle-layer training-set subjects was calculated between the two groups.
 - c. The resulting p -values were sorted in ascending order.
 - d. The feature vector was initiated by taking the feature with the smallest p -value, thus, the initial length was one.
 - e. For this feature vector, the classification accuracy was calculated with 10-fold cross-validation, thus, the middle-layer training set was divided into an inner-layer 10-fold partition with an inner-layer training- and testing set.
 - f. The next feature from the sorted list was added. For this feature vector, the inner-layer classification with 10-fold out cross-validation was repeated.
 - g. The result was compared to the previous result. The added entry to the feature vector was included only if the following three criteria were met:
 - The resulting classification accuracy was required to be at least as high as the maximum of the previously obtained classification accuracies; that is, the second accuracy had to be larger than the first entry, or the 6th accuracy had to be larger than the previous 5 classification accuracies.
 - If the so far best sensitivity/specificity, or in other words, accuracy for members of the first group/second group, respectively, was lower than 0.75, then the obtained sensitivity had to be at least as large as this maximum.
 - If the so far best specificity/sensitivity, was lower than 0.5, then the obtained specificity had to be larger than this maximum.
 - h. This way, features were added and tested for their contribution to the classification accuracy until all available features were used, or until the feature vector reached a

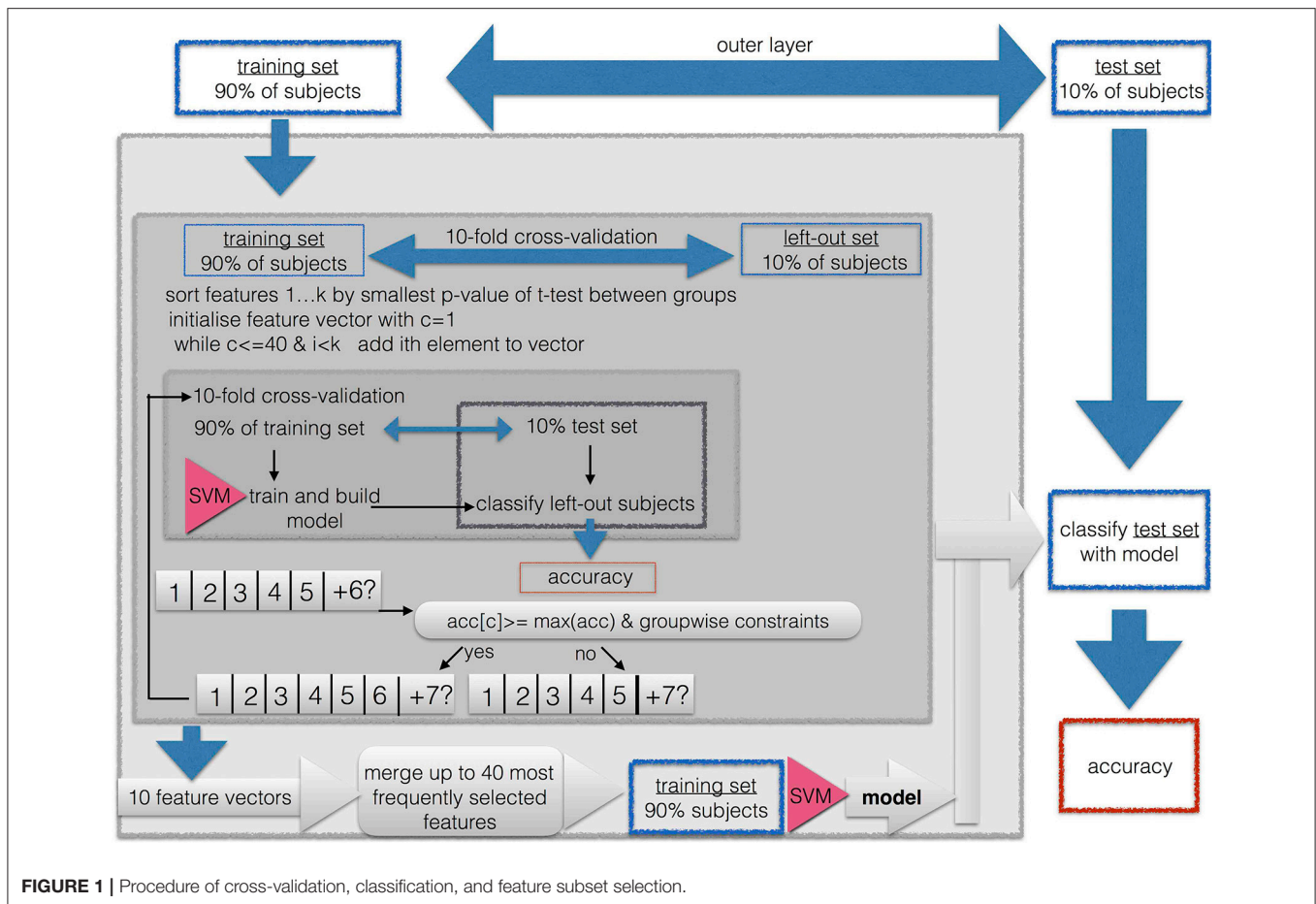


FIGURE 1 | Procedure of cross-validation, classification, and feature subset selection.

maximum of 40 entries, or if more than a consecutive number of 10% of all available features was not added to the feature vector.

- The resulting 10 optimized feature sets were concatenated and the occurrences of the features were counted. A final feature vector was formed by including only those features which were selected at least in 3 of the 10 iterations. If this resulted in no features, all features were included that were selected at least in 2 out of 10 iterations, and if this still did not yield any feature, this threshold was lowered to 1 out of 10 iterations. If the resulting feature vector included more than 41 features, only the top-most selected 41 features were included (equaling to approximately 90% of the available SPECT features).
- The resulting feature vector was used to train a SVM on the outer-layer training set, and the resulting model was used to classify the outer-layer test set, which was then used to calculate the general classification accuracy and the within-group accuracy for the two subgroups (i.e., sensitivity/specificity).

The thresholds of 0.75 and 0.5 were selected as rough estimator for above-chance classification; a value of 0.75 can be considered to be clearly above chance, while values below 0.5 are considered

to be clearly below chance and thus, a result of overfitting the model to one of the two groups.

Feature subset selection and classification was done for each of the scenarios as described in Section 2.8 and separately for each of the 6 combinations of groups.

2.11. Statistics

We calculated overall group classification accuracy, but also accuracies for the single groups, that can be understood in a sense of sensitivity and specificity. For sensitivity and specificity we have to define what are the positives and what the negatives, which is not directly applicable to pairwise group classifications. Thus, the accuracy of the single groups was

$$acc_{group} = \frac{N \text{ correct in group}}{\text{total } N \text{ of group}} \quad (1)$$

Namely, for each group the proportion of correctly classified individuals was determined in each of the classification situations (feature vectors and group combinations).

In order to evaluate the resulting accuracies we calculated the maximum-chance criterion, that is the proportion of samples contained in the larger of the two groups of one group comparison.

Wilcoxon-tests, *t*-tests, or Fisher's exact tests were used as appropriate for pairwise group-comparisons of numerical or nominal data characteristics of the samples. We applied Bonferroni correction to the resulting *p*-values by interpreting them at the level $0.05/(16 \times 6)$ for the 6 group comparisons and 16 neuropsychological scales and demographic aspects.

3. RESULTS

3.1. Sample Details

The demographic details as well as the results of the neuropsychological scales of the patients are given group-wise in **Table 1**.

The results of the pairwise group comparisons are shown in **Table 2**.

3.2. Classification Results

The results of the classification are given in **Table 3**. We marked the best classification accuracies in bold font, where the best accuracy was defined as the highest overall accuracy and also high within group accuracies. We can see that for all comparisons involving AD, that is, aSCC-AD, aMCI-AD, and AD-DCI, the best result was obtained when combining a single EEG measure with SPECT. For the comparisons of aSCC-aMCI and aMCI-DCI the best comparison was obtained when merging all EEG measures, and adding SPECT to this configuration yielded the same result. For the comparison aSCC-DCI the best result was found for single EEG measures, but the accuracies were below the maximum-chance criterion.

Please note that, however, the combination leads to a re-ordering of the features during the sorting according to *p*-values, so that merging of EEG and SPECT does not necessarily mean that there were actually features from the EEG or SPECT included in the analysis. For example, the classification accuracy for aSCC vs. AD was already quite high when using SPECT alone. When introducing the EEG measure spectrum, none of the EEG measures was finally used, but the additional features in the feature vector helped to choose the most informative SPECT values so that the accuracy was higher.

3.3. Visualization of Group Differences

For clinical interpretability we created heatmaps for all measures. Since EEG+SPECT yielded most informative measures, we based this illustration on the features selected from this combination. The heatmaps represent *t*-values for the pairwise group comparisons of the EEG, where all non-used indices were set to zero. Thus, we highlighted the region-interactions/frequencies that were selected during feature subset selection. In addition, we noted which SPECT regions were included into the analysis.

We include here only two measures as examples, while the others can be retrieved in the Supplementary Material Section. We include transfer function (*h*) which is the base on which the other measures are calculated, and which indeed yields reasonable accuracies for several comparisons.

From **Figure 2** we can see that from the transfer function polynomial, single channels are selected because of the information spread from these channels toward others. For

most comparisons, the classifier was based on at least one such interaction where the strength was higher in the one than in the other group and at least one such interaction with the opposite pattern.

In contrast, in **Figure 3** we can see for real valued coherence the typical pattern of information contained in the lower frequencies, where patients with aSCC showed higher values than the other groups, followed by DCI and then AD. Single electrode interactions were chosen, and most information was contained in the frequency ranges delta, theta and alpha, while beta contributed only with a single value for AD vs. DCI and the gamma range was not informative, at all.

The regions typically used from SPECT (**Figure 4**) are quite consistent across the EEG measures, especially for the comparisons with the AD group. Patients with AD have lower perfusion values in bilateral parietotemporal cortex, medial, lateral, and posterior temporal-lobe, and the temporal pole. In addition, differences in the cerebellum (cortex and white matter), the occipital cortex, and the thalamus were useful sources for information. However, while all the regions mentioned here were found to show lower perfusion values in AD than all other groups, the cerebellar white matter evokes higher values in AD compared to DCI.

4. DISCUSSION

In this work, we examined the diagnostic accuracy of quantitative EEG and SPECT alone and in combination with each other in order to differentiate patients with AD, aMCI, aSCC, and DCI. SCC are common in the elderly population and can be an early phase of MCI (Kryscio et al., 2014). Patients with SCC are twice as likely to develop AD than people without SCC (Mitchell et al., 2014). Conversion rates of MCI to AD are estimated around 10–18% per year (Gauthier et al., 2006), 11–33% after 2 years (Ritchie, 2004), and 50–70% after 3–5 years (see review in Rossini et al., 2007). Depressive symptoms in the elderly affect daily living and severely reduce quality of life (Stögmänn et al., 2016a). Depressive symptoms correlate with conversion from MCI to AD (Makizako et al., 2016; Stögmänn et al., 2016b), and can challenge differential diagnosis (Leyhe et al., 2017). Early differential diagnosis between these disorders with amnesic symptoms is a prerequisite to targeted interventions.

We found that for specific comparisons, a combination of EEG and SPECT yields the best diagnostic accuracy, while for other group contrasts, the one or the other modality is superior. In the following, we want to discuss our results in relation to previously reported classification approaches and we want to emphasize the novelty of a possible classification of DCI by using EEG and SPECT in combination.

4.1. EEG—An Underestimated Source of Information?

Previous research has suggested that biomarkers from the EEG may be more useful than methods investigating cerebral perfusion, such as HMPAO SPECT in order to identify patients suffering from AD at an early stage of the condition (Gungor

TABLE 1 | Sample overview with means and standard deviations in parentheses for neuropsychological test results.

| Sample | aSCC | aMCI | AD | DCI |
|--|---------------|---------------|---------------|--------------|
| <i>N</i> | 41 | 71 | 39 | 69 |
| Median age | 68 | 70 | 76 | 69 |
| Mean age | 67.54 | 70.04 | 74.97 | 68.81 |
| Age range | 52–82 | 51–87 | 57–90 | 50–86 |
| <i>N</i> women | 30 | 38 | 27 | 51 |
| HADS | | | | |
| Anxiety (SD) | 52.46 (10.30) | 50.72 (11.61) | 50.91 (11.27) | 61.45 (9.21) |
| Depression (SD) | 53.54 (5.66) | 53.68 (8.82) | 55.15 (8.85) | 71.18 (6.60) |
| CERAD | | | | |
| Semantic verbal fluency <i>z</i> -value (SD) | 0.30 (0.90) | −0.60 (0.90) | −1.73 (0.86) | −0.50 (1.15) |
| Boston naming test <i>z</i> -value (SD) | 0.49 (0.80) | −0.26 (1.06) | −1.09 (1.50) | −0.21 (1.35) |
| MMSE raw (SD) | 28.66 (1.06) | 26.80 (1.66) | 20.23 (4.50) | 27.81 (1.62) |
| Wordlist learning <i>z</i> -value (SD) | −0.51 (0.97) | −1.67 (0.89) | −3.31 (1.28) | −1.03 (1.24) |
| Wordlist recall <i>z</i> -value (SD) | −0.37 (0.64) | −1.60 (0.91) | −2.78 (0.93) | −0.77 (0.98) |
| Wordlist recognition <i>z</i> -value (SD) | −0.10 (0.91) | −1.46 (1.30) | −2.56 (1.41) | −0.40 (1.41) |
| Figures copying <i>z</i> -value (SD) | 0.50 (0.94) | 0.13 (1.14) | −1.29 (1.95) | 0.50 (1.18) |
| Figures recall <i>z</i> -value (SD) | 0.55 (1.19) | −1.08 (1.33) | −2.16 (1.02) | 0.32 (1.53) |
| Figures recognition <i>z</i> -value (SD) | 0.36 (1.03) | −0.90 (1.21) | −1.69 (1.23) | 0.03 (1.15) |
| PLUS-TESTS | | | | |
| Trail making test A <i>z</i> -value (SD) | 0.81 (1.06) | −0.16 (1.18) | −1.73 (1.28) | −0.04 (1.54) |
| Trail making test B <i>z</i> -value (SD) | 0.48 (1.27) | −0.32 (1.11) | −1.25 (0.83) | 0.31 (1.21) |
| Phonematic verbal fluency <i>z</i> -value (SD) | 0.44 (0.99) | 0.08 (1.07) | −0.55 (1.15) | −0.05 (1.25) |

N, number; aSCC, amnesic subjective cognitive complaints; aMCI, mild cognitive impairment.

AD, Alzheimer's disease; DCI, depression with cognitive impairment; SD, standard deviation.

z-values of the CERAD scores refer to the relative scores with respect to a normative (cognitively healthy) group and adjusted for age and education.

TABLE 2 | Sample comparisons test-value/*p*-values.

| Comparison | aSCC-aMCI | aSCC-AD | aSCC-DCI | aMCI-AD | aMCI-DCI | AD-DCI |
|-----------------------------|------------|-------------|------------|-------------|-------------|-------------|
| Age (<i>t</i> -test) | −1.55/0.12 | −4.15/* | −0.81/0.42 | −2.79/0.006 | 0.84/0.40 | 3.56/<0.001 |
| Sex (Fisher) | 2.37/0.05 | 1.21/0.81 | 0.96/1 | 0.51/0.16 | 0.41/0.01 | 0.79/0.66 |
| HADS | | | | | | |
| Anxiety (Wilc) | 0.78/0.43 | 0.94/0.35 | −4.52/* | 0.14/0.89 | −5.48/* | −4.76/* |
| Depression (Wilc) | 0.26/0.80 | −1.02/0.31 | −8.39/* | −1.11/0.27 | −8.80/* | −7.19/* |
| CERAD | | | | | | |
| Semantic vf (Wilc) | 4.35/* | 6.98/* | 3.54/* | 5.36/* | −0.66/0.51 | −5.33/* |
| Boston naming test (Wilc) | 3.66/* | 4.90/* | 2.48/0.01 | 2.92/0.004 | −0.50/0.62 | −2.97/0.003 |
| MMSE (Wilc) | 5.55/* | 7.74/* | 2.54/0.01 | 8.17/* | −3.71/* | −8.37/* |
| Wordlist learning (Wilc) | 5.42/* | 7.15/* | 1.99/0.05 | 6.07/* | −3.65/* | −6.87/* |
| Wordlist recall (Wilc) | 6.76/* | 7.36/* | 1.87/0.06 | 5.48/* | −5.05/* | −7.31/* |
| Wordlist recognition (Wilc) | 5.52/* | 6.22/* | 1.14/0.26 | 3.91/* | −4.50/* | −5.89/* |
| Figures copying (Wilc) | 1.73/0.08 | 4.07/* | −0.90/0.37 | 3.74/* | −2.49/0.01 | −4.40/* |
| Figures recall (Wilc) | 5.48/* | 7.11/* | 0.72/0.47 | 4.35/* | −4.94/* | −6.80/* |
| Figures recognition (Wilc) | 5.04/* | 6.15/* | 1.28/0.20 | 3.52/* | −4.27/* | −6.00/* |
| PLUS-TESTS | | | | | | |
| Trail making test A (Wilc) | 3.92/* | 6.24/* | 2.87/0.004 | 5.03/* | −0.55/0.58 | −4.87/* |
| Trail making test B (Wilc) | 2.91/0.004 | 3.98/* | 0.39/0.70 | 2.66/0.008 | −2.74/0.006 | −3.65/* |
| Phonematic vf (Wilc) | 1.63/0.10 | 3.29/<0.001 | 1.78/0.07 | 2.34/0.02 | 0.21/0.83 | −2.08/0.04 |

vf, verbal fluency; Fisher's test, oddsRatio/*p*-value; *t*-test, *t*-value/*p*-value; Wilc: Wilcoxon, *z*-value/*p*-value; aSCC, amnesic subjective cognitive complaints; aMCI, mild cognitive impairment; AD, Alzheimer's disease; DCI, depression with cognitive impairment; *significant at Bonferroni-corrected level $p < 0.00052083$.

TABLE 3 | Sample comparison classification accuracies.

| | aSCC-aMCI | aSCC-AD | aSCC-DCI | aMCI-AD | aMCI-DCI | AD-DCI |
|---------------------------|------------------------|-------------------|-----------------------|-----------------------|-------------------|----------------|
| EEG SINGLE | | | | | | |
| DC | 0.36(0.5/0.29) | 0.6(0.6/0.6) | 0.6(0.67/0.57) | 0.5(0.5/0.5) | 0.5(0.4/0.6) | 0.6(0/1) |
| S | 0.64(0.75/0.57) | 0.8(0.8/0.8) | 0.2(0.67/0) | 0.5(0.83/0) | 0.5(0.4/0.6) | 0.8(0.5/1) |
| h | 0.73(0.5/0.86) | 0.5(0.6/0.4) | 0.6(0.67/0.57) | 0.6(0.5/0.75) | 0.8(0.8/0.8) | 0.7(0.75/0.67) |
| Af | 0.55(0.25/0.71) | 0.3(0/0.6) | 0.2(0/0.29) | 0.6(0.5/0.75) | 0.7(0.8/0.6) | 0.5(0.75/0.33) |
| COH | 0.55(0.75/0.43) | 0.8(0.8/0.8) | 0.2(0.67/0) | 0.6(0.33/1) | 0.6(0.2/1) | 0.7(0.25/1) |
| iCOH | 0.36(0.5/0.29) | 0.5(0.6/0.4) | 0.5(0.33/0.57) | 0.5(0.5/0.5) | 0.7(0.4/1) | 0.5(0.25/0.67) |
| pCOH | 0.55(0.5/0.57) | 0.7(0.8/0.6) | 0.4(0.67/0.29) | 0.5(0.5/0.5) | 0.5(0.6/0.4) | 0.7(0.5/0.83) |
| PDC | 0.64(0.75/0.57) | 0.7(0.6/0.8) | 0.5(0.67/0.43) | 0.5(0.83/0) | 0.5(0.4/0.6) | 0.6(0.25/0.83) |
| PDCF | 0.27(0.25/0.29) | 0.8(1/0.6) | 0.6(0.33/0.71) | 0.4(0.33/0.5) | 0.6(0.4/0.8) | 0.6(0.75/0.5) |
| GPDC | 0.36(0.75/0.14) | 0.5(0.8/0.2) | 0.6(0.33/0.71) | 0.3(0.5/0) | 0.5(0.4/0.6) | 0.6(0.5/0.67) |
| DTF | 0.18(0.25/0.14) | 0.5(0.8/0.2) | 0.3(0.67/0.14) | 0.5(0.67/0.25) | 0.8(0.8/0.8) | 0.5(0.5/0.5) |
| dDTF | 0.27(0.25/0.29) | 0.5(0.4/0.6) | 0.5(0.67/0.43) | 0.5(0.83/0) | 0.8(1/0.6) | 0.5(0.25/0.67) |
| ffDTF | 0.55(0.75/0.43) | 0.4(0.4/0.4) | 0.4(0.67/0.29) | 0.4(0.5/0.25) | 0.6(0.2/1) | 0.6(0.25/0.83) |
| GGC | 0.36(0.5/0.29) | 0.6(0.4/0.8) | 0.6(0.33/0.71) | 0.4(0.33/0.5) | 0.5(0.2/0.8) | 0.6(0.25/0.83) |
| SPECT | 0.18(0/0.29) | 0.8(1/0.6) | 0.4(0.33/0.43) | 0.5(0.33/0.75) | 0.5(0.6/0.4) | 0.7(0.75/0.67) |
| EEG SINGLE + SPECT | | | | | | |
| DC + SPECT | 0.45(0.25/0.57) | 0.9(1/0.8) | 0.5(0.33/0.57) | 0.6(0.5/0.75) | 0.5(0.4/0.6) | 0.7(0.75/0.67) |
| S + SPECT | 0.64(0.75/0.57) | 0.9(1/0.8) | 0.3(0/0.43) | 0.5(0.33/0.75) | 0.5(0.4/0.6) | 0.8(0.75/0.83) |
| h + SPECT | 0.73(0.5/0.86) | 0.8(1/0.6) | 0.6(0/0.86) | 0.7(0.83/0.5) | 0.8(0.8/0.8) | 1(1/1) |
| Af + SPECT | 0.55(0.25/0.71) | 0.8(0.8/0.8) | 4(0/0.57) | 0.7(0.67/0.75) | 0.7(0.8/0.6) | 0.7(0.75/0.67) |
| COH + SPECT | 0.55(0.75/0.43) | 0.9(1/0.8) | 0.6(0.33/0.71) | 0.6(0.5/0.75) | 0.6(0.2/1) | 0.6(0.5/0.67) |
| iCOH + SPECT | 0.55(0.5/0.57) | 0.9(1/0.8) | 0.5(0/0.71) | 0.5(0.5/0.5) | 0.8(0.6/1) | 0.7(0.5/0.83) |
| pCOH + SPECT | 0.55(0.5/0.57) | 0.7(0.8/0.6) | 0.5(0.67/0.43) | 0.6(0.5/0.75) | 0.5(0.6/0.4) | 0.9(0.75/1) |
| PDC + SPECT | 0.64(0.75/0.57) | 0.7(1/0.4) | 0.4(0/0.57) | 0.5(0.33/0.75) | 0.5(0.4/0.6) | 0.7(0.5/0.83) |
| PDCF + SPECT | 0.27(0.25/0.29) | 0.9(1/0.8) | 0.4(0/0.57) | 0.5(0.5/0.5) | 0.6(0.4/0.8) | 0.7(0.75/0.67) |
| GPDC + SPECT | 0.36(0.75/0.14) | 0.9(1/0.8) | 0.4(0/0.57) | 0.6(0.5/0.75) | 0.5(0.4/0.6) | 0.7(0.75/0.67) |
| DTF + SPECT | 0.18(0.25/0.14) | 0.7(0.8/0.6) | 0.4(0.67/0.29) | 0.5(0.5/0.5) | 0.8(0.8/0.8) | 0.8(0.75/0.83) |
| dDTF + SPECT | 0.27(0.25/0.29) | 0.6(0.4/0.8) | 0.5(0.67/0.43) | 0.4(0.33/0.5) | 0.8(1/0.6) | 0.8(0.5/1) |
| ffDTF + SPECT | 0.55(0.75/0.43) | 0.7(0.8/0.6) | 0.5(0.33/0.57) | 0.5(0.33/0.75) | 0.6(0.2/1) | 0.6(0.25/0.83) |
| GGC + SPECT | 0.36(0.5/0.29) | 0.9(0.8/1) | 0.6(0.33/0.71) | 0.4(0.33/0.5) | 0.5(0.2/0.8) | 0.7(0.25/1) |
| EEG merged | 0.82(0.75/0.86) | 0.8(0.6/1) | 0.6(0.33/0.71) | 0.4(0.67/0) | 0.9(1/0.8) | 0.6(0.25/0.83) |
| EEG merged + SPECT | 0.82(0.75/0.86) | 0.7(0.4/1) | 0.6(0.33/0.71) | 0.5(0.67/0.25) | 0.9(1/0.8) | 0.6(0.25/0.83) |
| EEG graph | 0.55(0.75/0.43) | 0.7(0.4/1) | 0.2(0.33/0.14) | 0.5(0.67/0.25) | 0.4(0.2/0.6) | 0.7(0.5/0.83) |
| EEG graph + SPECT | 0.55(0.75/0.43) | 0.8(0.8/0.8) | 0.4(0.67/0.29) | 0.6(0.67/0.5) | 0.4(0.2/0.6) | 0.8(0.75/0.83) |
| Chance level | 0.63 | 0.65 | 0.63 | 0.65 | 0.51 | 0.64 |

aSCC, amnesic subjective cognitive complaints; aMCI, amnesic mild cognitive impairment; AD, Alzheimer's disease; DCI, depression with cognitive impairment; bold font, best result; S, auto- and cross-spectrum; DC, direct causality; h, transfer function; Af, transfer function polynomial; COH, real valued coherence; iCOH, complex coherence; pCOH, partial coherence; PDC, partial directed coherence; PDCF, partial directed coherence factor; GPDC, generalized partial directed coherence; DTF, directed transfer function; dDTF, direct directed transfer function; ffDTF, full frequency directed transfer function; GGC, Geweke's Granger causality; chance level, maximum-chance criterion according to maximum of group proportions.

et al., 2005). Still, the additional contribution of EEG seems to be underestimated, since EEG alterations such as slow theta-delta activity are a common feature of dementia and natural aging, as well (Rossini et al., 2007). In our study, alterations in the delta, theta, and alpha frequency ranges were prominent when comparing patients with AD to the other groups in widespread regions, where the exact localization of the most informative region depended highly on the measure of interaction. In patients with AD, coherence was lower than in patients with SCC, and than in patients with DCI, while patients with aMCI showed lower coherence than patients with AD. Previous research

reported that within and between hemisphere alpha coherence values are reduced in patients with dementia that show abnormal regional cerebral blood flow (Sloan et al., 1994). We could extend this finding by showing directly that combination of measures of interaction, for example partial coherence, with SPECT provides considerable information gain in a differential diagnostic setting. However, our results also demonstrate that the clear findings reported in the literature depend highly on the choice of the measure.

We want to mention that we performed a rather simple feature merging algorithm, and also the feature subset selection

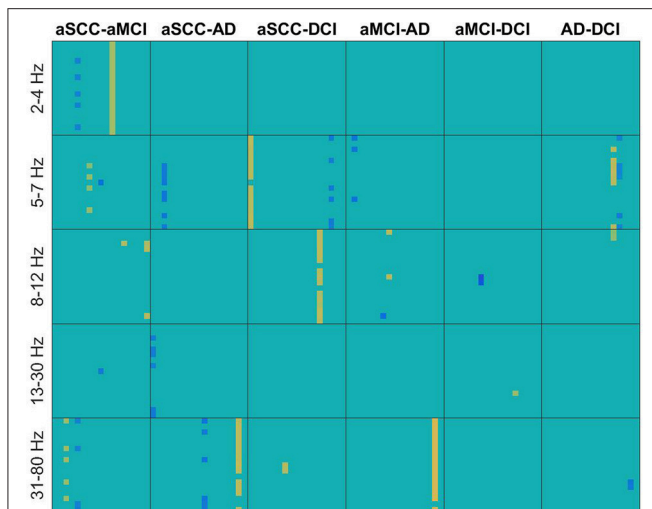


FIGURE 2 | Heatmaps of the t -values of group differences of all electrode \times electrode interactions for transfer function, sorted by groups comparisons in columns and frequency ranges in rows. Colors indicate values from -4.11 (dark blue) over zero (green) to $+5.24$ (yellow). All values that were not included for classification were set to zero. If the first group of the group comparison (e.g., aSCC in aSCC-aMCI) has higher values than the second group, this results in a positive t -value, i.e., yellow colors. Electrodes start from top to bottom and from left to right following the order: F3, F4, C3, C4, P3, P4, O1, O2, F7, F8, T3, T4, T5, T6, Fz, Cz, and Pz. AD, Alzheimer's disease; DCI, depression with cognitive impairment; aMCI, mild cognitive impairment with amnesic symptoms; aSCC, subjective cognitive complaints with amnesic symptoms.

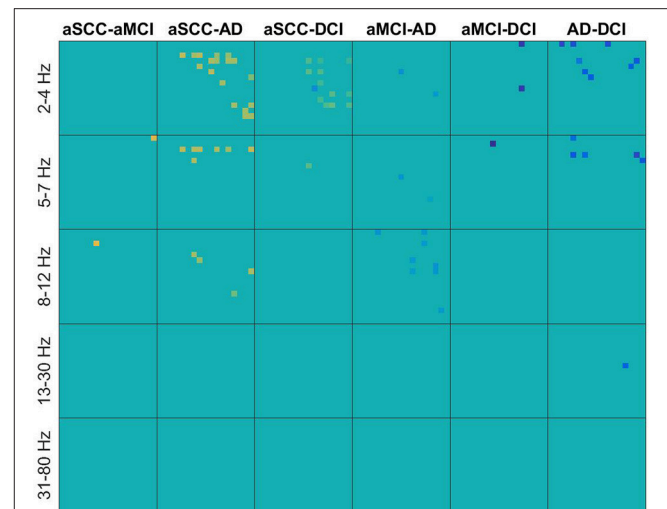


FIGURE 3 | Heatmaps of the t -values of group differences of all electrode \times electrode interactions for real valued coherence, sorted by groups comparisons in columns and frequency ranges in rows. Colors indicate values from -4.11 (dark blue) over zero (green) to $+5.24$ (yellow). All values that were not included for classification were set to zero. If the first group of the group comparison (e.g., aSCC in aSCC-aMCI) has higher values than the second group, this results in a positive t -value, i.e., yellow colors. Electrodes start from top to bottom and from left to right following the order: F3, F4, C3, C4, P3, P4, O1, O2, F7, F8, T3, T4, T5, T6, Fz, Cz, and Pz. AD, Alzheimer's disease; DCI, depression with cognitive impairment; aMCI, mild cognitive impairment with amnesic symptoms; aSCC, subjective cognitive complaints with amnesic symptoms.

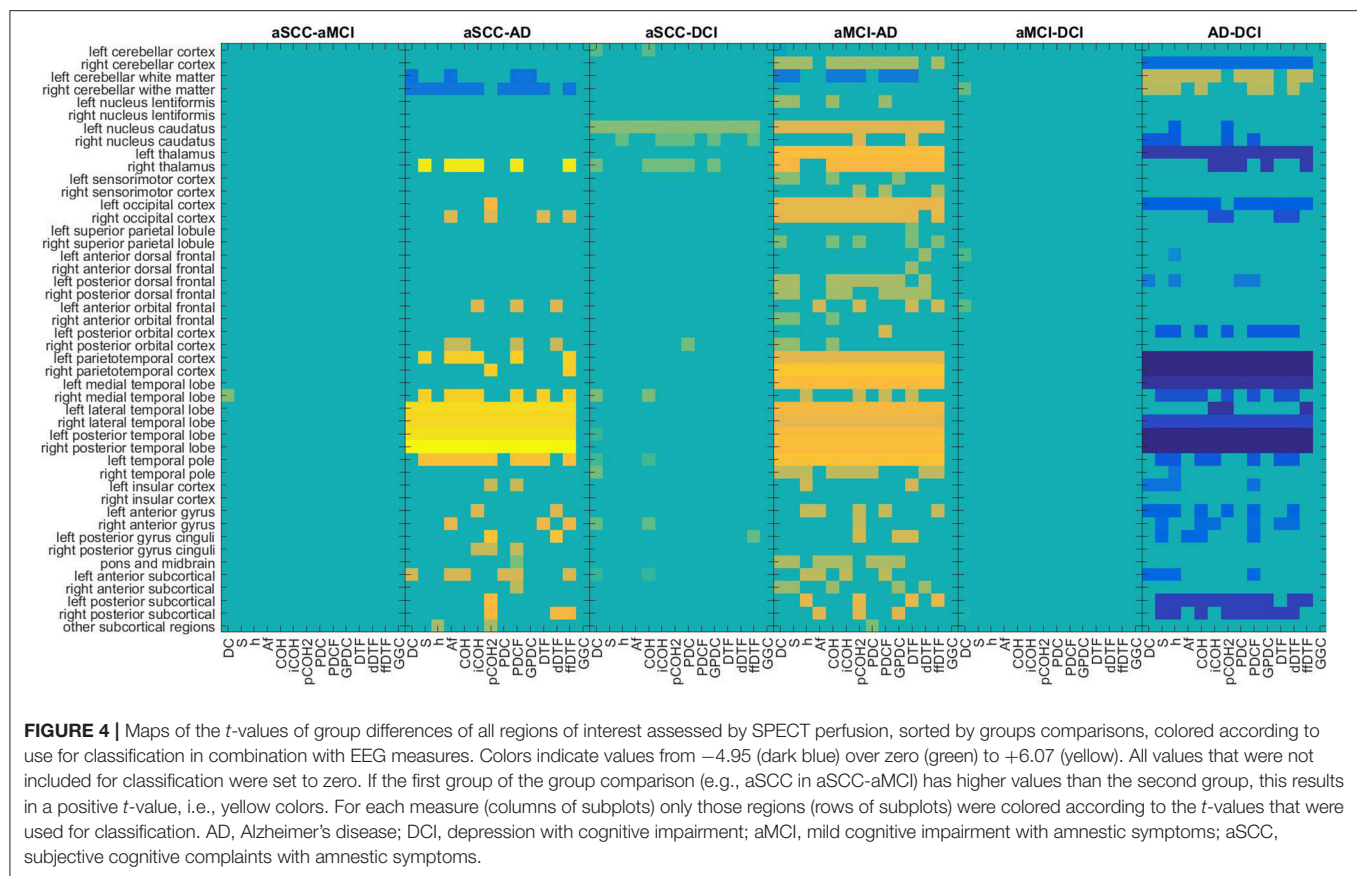
technique presented here is not able to fully explore the information in the data. In order to reduce computational complexity, the feature vectors were sorted by p -values. Processing the features in a different order might have yielded different results, which is also emphasized by the case when adding SPECT to EEG spectrum changes the results, even when the information from SPECT might not have been used (as found for spectrum). With more sophisticated feature subset selection techniques and feature merging algorithms we might achieve even higher accuracies.

The largest difference between information content in EEG and SPECT is seen for aSCC vs. DCI, where the best result is obtained with EEG-measures, only. However, the resulting accuracies are at chance, so that it is likely that none of the two modalities is able to accurately differentiate these two disorders. In contrast, the comparison of aSCC vs. aMCI and aMCI vs. DCI was highest when the best features from all EEG measures were merged, where this result did not change when including SPECT to the feature vector. The evidence for SPECT being useful to identify SCC or aSCC is scarce (Banzo et al., 2011; Frisoni et al., 2014). The differential diagnosis of aSCC is a challenge. In our study, we included patients with minimal deviations on the neuropsychological scales for memory, but who did not yet fulfill the clinical criteria for aMCI. Nevertheless, whether aSCC is a state of normal aging, where the patients become aware of the natural decay of memory capacities, or whether this is the first sign of a beginning dementia cannot be determined by neuropsychological scales, unless one has longitudinal data at his

disposal. The group in our study may be very heterogeneous, for these reasons. On this background it is remarkable that we were able to report above-chance classification accuracies of the EEG biomarkers.

EEG also successfully differentiated DCI from aMCI, best when merging all EEG measures, and from AD, combination with SPECT, yielding reasonable classification accuracies. Only the comparison of aSCC vs. DCI was not above chance with none of the applied feature vectors. A similar classification experiment of DCI vs. AD, aMCI and aSCC was—to our best knowledge—never done before with EEG, so that this result points to a new field of application. Especially in aMCI or AD depression is not rare and the differential diagnosis is often based on the trend of the symptoms when treating the depression adequately. Cognitive improvement after antidepressive therapy suggests that the depression, not a neurodegenerative disorder, causes the symptoms. As a conclusion the diagnosis of DCI can be made. However, since dementia and depressive symptoms coexist in some cases it could be difficult to assess whether depression is the cause or the effect of the cognitive impairment and vice versa. This is especially true when considering that depression is suspected to play a role in the progression of aMCI to AD (Van der Musselle et al., 2014; Chung et al., 2016).

Using robust invariant features from unprocessed EEGs, it may even be possible to reach higher classification accuracies than in the present manuscript (Buscema et al., 2015; Dimitriadis et al., 2015). However, in our study we used strict nested



cross-validation, which is the state of the art in order to avoid overfitting during parameter selection, and could rely on our sample with a sufficient size without need for data augmentation techniques as implemented in other studies (Dimitriadis et al., 2015). Moreover, the intention of this study was not to reach maximum classification accuracy of one particular method, but rather to show how EEG and SPECT could complement each other, while trying to render the comparison between individual and combined methods as fair as possible. However, our results are comparable with previous publications (Buscema et al., 2015; Gallego-Jutgla et al., 2015; Hatz et al., 2015). Other studies using entropy measures instead of measures of interaction report results with accuracies of 91.7–93.8% when discriminating MCI, AD and normal controls (McBride et al., 2015). After all, there was no healthy control group in our study, and the comparison to healthy controls is more straightforward and clinically not of interest, because differential diagnosis between AD and healthy or even aSCI can be accomplished reliably with classical paper and pencil tests. In contrast, we examined also the more challenging and interesting discrimination of DCI vs. AD or vs. aMCI yielding excellent classification accuracies.

4.2. Information Gain or Information Loss through Graph-Theory

It was suggested that graph-theoretical approaches could help to make measures of interaction more useful for the prediction

of MCI progression from the EEG (Vecchio et al., 2014, 2015; Miraglia et al., 2016; Rossini et al., 2016; Vecchio et al., 2016). In our study, using the measures of interaction directly yielded higher accuracies than the use of the derived graph-theoretic indices. Only for aSCC vs. AD and for AD vs. DCI above-chance classification (0.8) was obtained with graph-theoretical measures. This means that the way the information is integrated with graph-theoretical measures may not be advantageous in every scenario and needs to be examined from case to case.

4.3. HMPAO-SPECT

A systematic review found sensitivity and specificity of HMPAO-SPECT to distinguish AD from healthy controls to be 76.1 and 85.4%, respectively, and the distinction of vascular dementia and dementia with Lewy Bodies from AD yielded even lower diagnostic values (Yeo et al., 2013). We want to emphasize that when contrasting HMPAO-SPECT of AD and healthy controls, sensitivities and specificities are high: 81 and 96% (Fleming et al., 2002), or 91 and 86% (Johnson et al., 1993). However, when cases with diagnostic uncertainty are examined, only very low values with a sensitivity of 71–77% and a specificity of 38–44% can be achieved (Doran et al., 2005). It is also hard to identify AD among unselected patients in a memory clinic, resulting in a sensitivity of 75% and a specificity of 52% (Masterman et al., 1997). Indeed, a systematic review found that the diagnostic accuracy of HMPAO-SPECT to discriminate between AD and other forms of dementia

was characterized by a sensitivity of 71.3% and a specificity of 75.9% (Dougall et al., 2004a). This is also reflected by our results, where the highest accuracy values when using the SPECT-feature vector, only, were found for aSCC vs. AD. In clinical terms, this is the most obvious differentiation, followed by the more challenging contrasts of AD vs. DCI and then by AD vs. aMCI. There is a statistically significant difference in perfusion in specific brain areas between AD and aMCI (Fröhlich et al., 1989; Staffen et al., 2006, 2009; Tranfaglia et al., 2009; Van Heertum et al., 2009; Farid et al., 2011), but according to our results, it is not enough for creating a model with high distinctiveness when being used without further information, such as the EEG. It is worth to stress once again that our results were based on a quantitative evaluation, while many of the diagnostic characteristics of SPECT are based on expert ratings. The sensitivity of these ratings was found to be negatively correlated with the importance the expert attributed to regional hypoperfusion in the parietal lobes (Dougall et al., 2004b).

The contribution of HMPAO SPECT to the differentiation of DCI and other forms of cognitive impairment is well in line with the finding, that depression and specifically treatment-resistant depression shows significant alterations in circumscribed brain regions such as the hippocampus and the amygdala (Bonne et al., 1996; Mozley et al., 1996; Hornig et al., 1997; Kowatch et al., 1999; Cho et al., 2002). In patients with AD and depression, a selective hypoperfusion in the anterior and posterior cingulate gyrus and in the precuneus was reported (Liao et al., 2003). A direct comparison between patients with AD and DCI showed differences in perfusion in the left parieto-occipital lobe (Stoppe et al., 1995). Thus, it is likely that the contribution of SPECT to EEG can be explained by complementary information about regional abnormalities in DCI that differ from those of AD. Indeed, the regions that differ between aMCI and AD are also most informative when differentiating AD from DCI. Future work should have a closer look at the distinctive characteristics of DCI, where only a narrow range of publications have identified promising biomarkers.

4.4. Limitations

Firstly, this retrospective study cannot indicate which markers are important for prognostic questions. Nevertheless, prognosis is the most important question in this patient population. Therefore, future studies with longitudinal, prospective design are needed to clarify the role of EEG and SPECT in these respects.

Secondly, the ground truth of our sample is based on multimodal clinical assessment. That is, we have no post-mortem determination of definitive AD. This implies that the ground truth is somewhat unclear and that the diagnoses that were used for classification are not all correct. In addition, this means that SPECT and sometimes also EEG were part of the basis on which the clinician defined the diagnosis, which is in turn, our ground truth. This is the typical scenario in the clinics, but still, a drawback of retrospective studies. However, as described in Section 2.2, the EEG examination was not used to define one of the examined diagnoses, but to disclose epilepsy or other disorders that could cause the amnesic symptoms. Similarly, SPECT was only included in the diagnostic process

for differential diagnosis of disorders that were not included in the presented analysis. Moreover, the examination of EEG and SPECT at the time of diagnosis was performed only qualitatively, while the present work was based on quantitative analysis, only. In sum, we estimate the bias in our ground truth to be very small.

Third, the present study emphasizes that the EEG can be useful at the stage of aSCC. However, our study did not provide data from a healthy control group, mainly because it is difficult to obtain SPECT from healthy controls. Future studies using EEG will more easily recruit healthy controls and provide longitudinal data. The latter is important in order to demonstrate the predictive value of the identified biomarkers.

Fourth, we could not report the medication history of the patients but we assume that only a minority of them were drug-naïve at the time of examination. Specifically antidepressants are commonly prescribed in the elderly and it is possible that they are prescribed more likely in the group of DCI, since these patients might have consulted the general practitioner before visiting the memory clinic.

Finally, there are other diagnostic modalities such as structural MRI which show a very high diagnostic accuracy and increasing relevance in amnesic populations (Teipel et al., 2013). However, the purpose of this study was not to show the best method in order to contribute to the diagnosis, but to show whether the combination of EEG and SPECT is a valid approach. Especially EEG is a cheap and one of the most easily available diagnostic methods that could be integrated into the routine process of memory clinics.

5. CONCLUSIONS AND FUTURE DIRECTIONS

HMPAO SPECT alone cannot reliably identify AD and related disorders with memory problems, but its additive value in combination with other modalities is well acknowledged. Also the examination of the EEG has identified several useful biomarkers that could be considered for use in differential diagnosis of cognitive impairment in the elderly population.

Our data show that EEG outperforms SPECT in several differential diagnoses. We suggest that direct combination of these two modalities is very helpful since they are complementary to each other. Both EEG and SPECT are not the gold standard for the diagnosis of AD and aMCI; however, they are widely used and cost effective. Furthermore, EEG is a non-invasive investigation technique which can be administered many times during the course of the disease. It proved to be more discriminative even at the stage of aSCC. Combining SPECT with EEG should also be subject of further investigations, in order to technically optimize the diagnostic accuracy.

AUTHOR CONTRIBUTIONS

YH performed the analysis and wrote the first draft of the manuscript, which was revised by AB, RN, FR, ET, and WS. AB, AU, and AJ supervised the work in technical and statistical respects and contributed ideas to how the analysis should be

performed and how the results should be presented. NS extracted the EEG data, AL preprocessed the EEG data, JB extracted and pre-processed the SPECT data. MK and JB performed neuropsychological investigations. HZ supervised the work in neuropsychological respects. All of the listed authors have read, commented and approved the manuscript.

FUNDING

The presented research was funded by the Austrian Science Fund (FWF): KLIF 12-B00, I 2697-N31, and by the PMU FFF: A-11/02/004-TRI.

REFERENCES

- Aertsen, A., and Preissl, H. (1991). "Dynamics of activity and connectivity in physiological neuronal networks," in *Non Linear Dynamics and Neuronal Networks*, ed H. Schuster (New York, NY: VCH), 281–302.
- Babiloni, C., Ferri, R., Binetti, G., Cassarino, A., Forno, G. D., Ercolani, M., et al. (2006). Fronto-parietal coupling of brain rhythms in mild cognitive impairment: a multicentric eeg study. *Brain Res. Bull.* 69, 63–73. doi: 10.1016/j.brainresbull.2005.10.013
- Babiloni, C., Lizio, R., Marzano, N., Capotosto, P., Soricelli, A., Triggiani, A., et al. (2016). Brain neural synchronization and functional coupling in Alzheimer's disease as revealed by resting state EEG rhythms. *Int. J. Psychophysiol.* 103, 88–102. doi: 10.1016/j.ijpsycho.2015.02.008
- Banzo, I., Quirce, R., Martinez-Rodriguez, I., Jimnez-Bonilla, J., Portilla-Quattrociocchi, H., Medina-Quiroz, P., et al. (2011). Molecular neuroimaging in the study of cognitive impairment: Contribution of the cerebral blood flow SPECT with 99mTc-HMPAO and 18F-fdg pet/ct scan. *Rev. Espanola de Med. Nucl. (English Edition)* 30, 301–306. doi: 10.1016/j.remngl.2011.03.011
- Barnes, A., Lusman, D., Patterson, J., Brown, D., and Wyper, D. (2000). The use of statistical parametric mapping (SPM96) as a decision aid in the differential diagnosis of dementia using 99mTc-HMPAO SPECT. *Behav. Neurol.* 12, 77–86. doi: 10.1155/2000/482606
- Bonne, O., Krausz, Y., Gorfine, M., Karger, H., Gelfin, Y., Shapira, B., et al. (1996). Cerebral hypoperfusion in medication resistant, depressed patients assessed by tc99m HMPAO SPECT. *J. Affect. Disord.* 41, 163–171. doi: 10.1016/S0165-0327(96)00058-4
- Bonte, F., Hom, J., Tintner, R., and Weiner, M. (1990). Single photon tomography in Alzheimer's disease and the dementias. *Semin. Nucl. Med.* 20, 342–352. doi: 10.1016/S0001-2998(05)80238-7
- Brenner, R. (1999). "EEG and dementia," in *Electroencephalography. Basic Principles, Clinical Applications, and Related Fields*, 4th Edn., Chap. 19, eds E. Niedermeyer and F. Da Silva (Philadelphia, PA: Lippincott Williams & Wilkins), 349–359.
- Buscema, M., Vernieri, F., Massini, G., Scarscia, F., Breda, M., Rossini, P., et al. (2015). An improved I-FAST system for the diagnosis of Alzheimer's disease from unprocessed electroencephalograms by using robust invariant features. *Artif. Intell. Med.* 64, 59–74. doi: 10.1016/j.artmed.2015.03.003
- Cantero, J. L., Atienza, M., Cruz-Vadell, A., Suarez-Gonzalez, A., and Gil-Neciga, E. (2009). Increased synchronization and decreased neural complexity underlie thalamocortical oscillatory dynamics in mild cognitive impairment. *NeuroImage* 46, 938–948. doi: 10.1016/j.neuroimage.2009.03.018
- Chang, L. (1978). A method for attenuation correction in radionuclide computed tomography. *IEEE Trans. Nucl. Sci.* 21, 2–20. doi: 10.1109/TNS.1978.4329385
- Cho, M., Lyoo, I., Lee, D., Kwon, J., Lee, J., Lee, D., et al. (2002). Brain single photon emission computed tomography findings in depressive pseudodementia patients. *J. Affect. Disord.* 69, 159–66. doi: 10.1016/S0165-0327(01)00301-9
- Chung, J., Plitman, E., Nakajima, S., Chakravarty, M., Caravaggio, F., Takeuchi, H., et al. (2016). Depressive symptoms and small hippocampal volume accelerate the progression to dementia from mild cognitive impairment. *J. Alzheimers Dis.* 49, 743–754. doi: 10.3233/JAD-150679

ACKNOWLEDGMENTS

We thank the memory clinic and the EEG-staff for their routine work in collecting the data, Andreas Westiner for help in organizing the data, and Gert Dehnen for helpful suggestions for the figure design. Thanks to Omar Salah for final proofreading.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fnagi.2017.00290/full#supplementary-material>

- Dauwels, J., Vialatte, F., and Cichocki, A. (2010). Diagnosis of Alzheimer's disease from EEG signals: where are we standing? *Curr. Alzheimer Res.* 7, 487–505. doi: 10.2174/156720510792231720
- DeKosky, S., and Marek, K. (2003). Looking backward to move forward: early detection of neurodegenerative disorders. *Science* 302, 830–834. doi: 10.1126/science.1090349
- Dimitriadis, S., Laskaris, N., Bitzidou, M., Tarnanas, I., and Tsolaki, M. (2015). A novel biomarker of amnesic MCI based on dynamic cross-frequency coupling patterns during cognitive brain responses. *Front. Neurosci.* 9:350. doi: 10.3389/fnins.2015.00350
- Doran, M., Vinjamuri, S., Collins, J., Parker, D., and Larner, A. (2005). Single-photon emission computed tomography perfusion imaging in the differential diagnosis of dementia: a retroSPECTive regional audit. *Int. J. Clin. Pract.* 59, 496–500. doi: 10.1111/j.1368-5031.2005.00499.x
- Dougall, N., Bruggink, S., and Ebmeier, K. (2004a). Systematic review of the diagnostic accuracy of 99mTc-HMPAO-SPECT in dementia. *Am. J. Geriatr. Psychiatry* 12, 554–570. doi: 10.1097/00019442-200411000-00002
- Dougall, N., Nobili, F., and Ebmeier, K. (2004b). Predicting the accuracy of a diagnosis of Alzheimer's disease with 99mTc HMPAO single photon emission computed tomography. *Psychiatry Res.* 131, 157–168. doi: 10.1016/j.psychres.2003.11.001
- Farid, K., Caillat-Vigneron, N., and Sibon, I. (2011). Is brain SPECT useful in degenerative dementia diagnosis? *J. Comput. Assist. Tomogr.* 35, 1–3. doi: 10.1097/RCT.0b013e3181f56fda
- Fleming, J., Kemp, P., Bolt, L., and Goatman, K. (2002). Measurement of cerebral perfusion volume and 99mTc-HMPAO uptake using SPECT in controls and patients with Alzheimer's disease. *Nucl. Med. Commun.* 23, 1057–1064. doi: 10.1097/00006231-200211000-00004
- Folstein, M., Folstein, S., and McHugh, P. (1975). "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *J. Psychiatr. Res.* 12, 189–198. doi: 10.1016/0022-3956(75)90026-6
- Fornito, A., Zalesky, A., and Breakspear, M. (2015). The connectomics of brain disorders. *Nat. Rev. Neurosci.* 16, 159–172. doi: 10.1038/nrn3901
- Frisoni, G., Bocchetta, M., Chetelat, G., Rabinovici, G., de Leon, M., Kaye, J., et al. (2014). Imaging markers for Alzheimer disease: which vs how. *Neurology* 81, 487–500. doi: 10.1212/WNL.0b013e31829d86e8
- Friston, K. (1995). Commentary and opinion: II. statistical parametric mapping: ontology and current issues. *J. Cereb. Blood Flow Metab.* 15, 361–370. doi: 10.1038/jcbfm.1995.45
- Fröhlich, L., Eilles, C., Ihl, R., Maurer, K., and Lanczik, M. (1989). Stage-dependent reductions of regional cerebral blood flow measured by HMPAO-SPECT in dementia of Alzheimer type. *Psychiatry Res.* 29, 347–350. doi: 10.1016/0165-1781(89)90085-1
- Fröhlich, L., Maurer, K., Eilles, C., Dierks, T., and Ihl, R. (1992). Diagnostic value of HMPAO-SPECT and topographic analysis of EEG in dementia of Alzheimer type. *Clin. Neuropharmacol.* 15, 192–193. doi: 10.1097/00002826-199201001-00101
- Gallego-Jutgla, E., Sole-Casals, J., Vialatte, F., Elgendi, M., Cichocki, A., and Dauwels, J. (2015). A hybrid feature selection approach for

- the early diagnosis of Alzheimer's disease. *J. Neural Eng.* 12:016018. doi: 10.1088/1741-2560/12/1/016018
- Gauthier, S., Reisberg, B., Zaudig, M., Petersen, R., Ritchie, K., Broich, K., et al. (2006). Mild cognitive impairment. *Lancet* 367, 1262–1270. doi: 10.1016/S0140-6736(06)68542-5
- Goldenberg, G., Podreka, I., Suess, E., and Deecke, L. (1989). The cerebral localization of neuropsychological impairment in Alzheimer's disease: a SPECT study. *J. Neurol.* 236, 131–138. doi: 10.1007/BF00314327
- Gueguen, B., Ancrì, D., Derouesne, C., Bourdel, M., Guillou, S., and Landre, E. (1991). Comparison of SPECT and quantified EEG features in Alzheimer's type dementia. *Neurophysiol. Clin.* 21, 377–387. doi: 10.1016/S0987-7053(05)80329-6
- Gungor, H., Yildiz, A., Aydin, F., Gungor, F., Boz, A., and Ozkaynak, S. (2005). Tc-99m HMPAO brain SPECT findings in mild and moderate Alzheimer's disease: correlation with event related potentials. *J. Neurol. Sci.* 234, 47–53. doi: 10.1016/j.jns.2005.03.034
- Hatz, F., Hardmeier, M., Benz, N., Ehrensperger, M., Gschwandtner, U., Rüegg, S., et al. (2015). Microstate connectivity alterations in patients with early Alzheimer's disease. *Alzheimer's Res. Ther.* 7:78. doi: 10.1186/s13195-015-0163-9
- Herrmann-Lingen, C., Buss, U., and Snaith, R. (2007). *HADS-D Hospital Anxiety and Depression Scale - Deutsche Version. Ein Fragebogen zur Erfassung von Angst und Depressivität in der Somatischen Medizin. Testdokumentation und Handanweisung, 2nd Edn.* Bern: Huber.
- Hornig, M., Mozley, D., and Amsterdam, J. (1997). HMPAO SPECT imaging in treatment-resistant depression. *Prog. Neuro Psychopharmacol. Biol. Psychiatry* 21, 1097–1114. doi: 10.1016/S0278-5846(97)00100-0
- Jelic, V., Johansson, S., Almkvist, O., Shigeta, M., Julin, P., Nordberg, A., et al. (2000). Quantitative electroencephalography in mild cognitive impairment: longitudinal changes and possible prediction of Alzheimer's disease. *Neurobiol. Aging* 21, 533–540. doi: 10.1016/S0197-4580(00)00153-6
- Johnson, K., Kijewski, M., Becker, J., Garada, B., Satlin, A., and Holman, B. (1993). Quantitative brain SPECT in Alzheimer's disease and normal aging. *J. Nucl. Med.* 34, 2044–2048.
- Kaneko, K., Kuwabara, Y., Sasaki, M., Ogomori, K., Ichimiya, A., Koga, H., et al. (2004). Posterior cingulate hypoperfusion in Alzheimer's disease, senile dementia of Alzheimer type, and other dementias evaluated by three-dimensional stereotactic surface projections using Tc-99m HMPAO SPECT. *Clin. Nucl. Med.* 29, 362–366. doi: 10.1097/01.rlu.0000127091.43180.92
- Kemp, P., Hoffmann, S., Holmes, C., Bolt, L., Ward, T., Holmes, R., et al. (2005). The contribution of statistical parametric mapping in the assessment of precuneal and medial temporal lobe perfusion by 99mTc-HMPAO SPECT in mild Alzheimer's and Lewy body dementia. *Nucl. Med. Commun.* 26, 1099–1106. doi: 10.1097/00006231-200512000-00009
- Klados, M., Styliadis, C., Frantzidis, C., Paraskevopoulos, E., and Bamidis, P. (2016). Beta-band functional connectivity is reorganized in mild cognitive impairment after combined computerized physical and cognitive training. *Front. Neurosci.* 10:55. doi: 10.3389/fnins.2016.00055
- Koenig, T., Prichep, L., Dierks, T., Hubl, D., Wahlund, L. O., John, E. R., et al. (2005). Decreased EEG synchronization in Alzheimer's disease and mild cognitive impairment. *Neurobiol. Aging* 26, 165–171. doi: 10.1016/j.neurobiolaging.2004.03.008
- Kowatch, R., Devous, M., Harvey, D., Mayes, T., Trivedi, M., Emslie, G., et al. (1999). A SPECT HMPAO study of regional cerebral blood flow in depressed adolescents and normal controls. *Prog. Neuro Psychopharmacol. Biol. Psychiatry* 23, 643–656. doi: 10.1016/S0278-5846(99)00023-8
- Krysic, R., Abner, E., Cooper, G., Fardo, D., Jicha, G., Nelson, P., et al. (2014). Self-reported memory complaints: implications from a longitudinal cohort with autopsies. *Neurology* 83, 1359–1365. doi: 10.1212/WNL.0000000000000856
- Kwa, V., Weinstein, H., Posthumus Meyjes, E., van Royen, E., Bour, L., Verhoeff, P., et al. (1993). SPECTral analysis of the EEG and 99m-Tc-HMPAO SPECT-scan in Alzheimer's disease. *Biol. Psychiatry* 33, 100–107. doi: 10.1016/0006-3223(93)90308-Z
- Ladeira, R., Diniz, B., Nunes, P., and Forlenza, O. (2009). Combining cognitive screening tests for the evaluation of mild cognitive impairment in the elderly. *Clinics (Sao Paulo)* 64, 967–973. doi: 10.1590/S1807-59322009001000006
- Lehnertz, K. (2011). Assessing directed interactions from neurophysiological signals - an overview. *Physiol. Meas.* 32, 1715–1724. doi: 10.1088/0967-3334/32/11/R01
- Leyhe, T., Reynolds, C., Melcher, T., Linnemann, C., Klöppel, S., Blennow, K., et al. (2017). A common challenge in older adults: classification, overlap, and therapy of depression and dementia. *Alzheimers Dement.* 13, 59–71. doi: 10.1016/j.jalz.2016.08.007
- Liao, Y., Liu, R., Lee, Y., Sun, C., Liu, C., Wang, P., et al. (2003). Selective hypoperfusion of anterior cingulate gyrus in depressed AD patients: a brain SPECT finding by statistical parametric mapping. *Dement. Geriatr. Cogn. Disord.* 16, 238–244. doi: 10.1159/000072808
- Lopez, O. (2013). Mild cognitive impairment. *Continuum (Minneapolis)* 19, 411–424. doi: 10.1212/01.CON.0000429175.29601.97
- Makizako, H., Shimada, H., Doi, T., Tsutsumimoto, K., Hotta, R., Nakakubo, S., et al. (2016). Comorbid mild cognitive impairment and depressive symptoms predict future dementia in community older adults: a 24-month follow-up longitudinal study. *J. Alzheimers Dis.* 54, 1473–1482. doi: 10.3233/JAD-160244
- Marple, S. (1987). *Digital SPECTral Analysis with Applications*. Englewood Cliffs, NJ: Prentice Hall.
- Masterman, D., Mendez, M., Fairbanks, L., and Cummings, J. (1997). Sensitivity, specificity, and positive predictive value of technetium 99-HMPAO SPECT in discriminating Alzheimer's disease from other dementias. *J. Geriatr. Psychiatry Neurol.* 10, 15–21. doi: 10.1177/089198879701000104
- McBride, J., Zhao, X., Munro, N., Jicha, G., Smith, C., and Jiang, Y. (2015). Discrimination of mild cognitive impairment and Alzheimer's disease using transfer entropy measures of scalp EEG. *J. Health Eng.* 6, 55–70. doi: 10.1260/2040-2295.6.1.55
- McKhann, G., Drachman, D., Folstein, M., Katzman, R., Price, D., and Stadlan, E. (1984). Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology* 34, 939–944. doi: 10.1212/WNL.34.7.939
- Miraglia, F., Vecchio, F., Bramanti, P., and Rossini, P. (2016). Eeg characteristics in “eyes-open” versus “eyes-closed” conditions: Small-world network architecture in healthy aging and age-related brain degeneration. *Clin. Neurophysiol.* 127, 1261–1268. doi: 10.1016/j.clinph.2015.07.040
- Mitchell, A., Beaumont, H., Ferguson, D., Yadegarfar, M., and Stubbs, B. (2014). Risk of dementia and mild cognitive impairment in older people with subjective memory complaints: meta-analysis. *Acta Psychiatr. Scand.* 130, 439–451. doi: 10.1111/acps.12336
- Montplaisir, J., Petit, D., McNamara, D., and Gauthier, S. (1996). Comparisons between SPECT and quantitative EEG measures of cortical impairment in mild to moderate Alzheimer's disease. *Eur. Neurol.* 36, 197–200. doi: 10.1159/000117247
- Morris, J., Heyman, A., Mohs, R., Hughes, J., van Belle, G., Fillenbaum, G., et al. (1989). The consortium to establish a registry for Alzheimer's disease (CERAD). Part I. clinical and neuropsychological assessment of Alzheimer's disease. *Neurology* 39, 1159–1165. doi: 10.1212/WNL.39.9.1159
- Mozley, P., Hornig-Rohan, M., Woda, A., Kim, H.-J., Alavi, A., Payer, F., et al. (1996). Cerebral HMPAO SPECT in patients with major depression and healthy volunteers. *Prog. Neuro Psychopharmacol. Biol. Psychiatry* 20, 443–458. doi: 10.1016/0006-3223(95)00421-1
- Müller, T., Thome, J., Chiamonti, R., Dierks, T., Maurer, K., Fallgatter, A., et al. (1997). A comparison of qEEG and HMPAO-SPECT in relation to the clinical severity of Alzheimer's disease. *Eur. Arch. Psychiatry Clin. Neurosci.* 247, 259–263. doi: 10.1007/BF02900304
- Petersen, R., Smith, G., Waring, S., Ivnik, R., Tangalos, E., and Kokmen, E. (1999). Mild cognitive impairment: clinical characterization and outcome. *Arch. Neurol.* 56, 303–308. doi: 10.1001/archneur.56.3.303
- Reitan, R. (1979). *Trail-Making Test*. Arizona: Reitan Neuropsychology Laboratory.
- Rentz, D., Rodriguez, M., Amariglio, R., Stern, Y., Sperling, R., and Ferris, S. (2013). Promising developments in neuropsychological approaches for the detection of preclinical Alzheimer's disease: a selective review. *Alzheimer's Res. Ther.* 5:58. doi: 10.1186/alzrt222
- Ritchie, K. (2004). Mild cognitive impairment: an epidemiological perspective. *Dialogues Clin. Neurosci.* 6, 401–408.
- Rodriguez, G., Nobili, F., Copello, F., Vitali, P., Gianelli, M., Taddei, G., et al. (1999). 99mTc-HMPAO regional cerebral blood flow and quantitative electroencephalography in Alzheimer's disease: a correlative study. *J. Nucl. Med.* 40, 522–529.

- Rossini, P., Di Iorio, R., Granata, G., Miraglia, F., and Vecchio, F. (2016). From mild cognitive impairment to Alzheimer's disease: a new perSPECTive in the "land" of human brain reactivity and connectivity. *J. Alzheimers Dis.* 53, 1389–1393. doi: 10.3233/JAD-160482
- Rossini, P., Rossi, S., Babiloni, C., and Polich, J. (2007). Clinical neurophysiology of aging brain: from normal aging to neurodegeneration. *Prog. Neurobiol.* 83, 375–400. doi: 10.1016/j.pneurobio.2007.07.010
- Rubinov, M., and Sporns, O. (2010). Complex network measures of brain connectivity: uses and interpretations. *Neuroimage* 52, 1059–1069. doi: 10.1016/j.neuroimage.2009.10.003
- Scheltens, P., Launer, L., Barkhof, F., Weinstein, H., and Jonker, C. (1997). The diagnostic value of magnetic resonance imaging and technetium 99m-HMPAO single-photon-emission computed tomography for the diagnosis of Alzheimer disease in a community-dwelling elderly population. *Alzheimer Dis. Assoc. Disord.* 11, 63–70. doi: 10.1097/00002093-199706000-00002
- Schlögl, A. (2006). A comparison of multivariate autoregressive estimators. *Signal Process.* 86, 2426–2429. doi: 10.1016/j.sigpro.2005.11.007
- Schlögl, A., and Brunner, C. (2008). BioSig: a free and open source software library for BCI research. *Computer* 41, 44–50. doi: 10.1109/MC.2008.407
- Sloan, E., Fenton, G., Kennedy, N., and MacLennan, J. (1994). Neurophysiology and SPECT cerebral blood flow patterns in dementia. *Electroencephalogr. Clin. Neurophysiol.* 91, 163–170. doi: 10.1016/0013-4694(94)90066-3
- Sloan, E., Fenton, G., Kennedy, N., and MacLennan, J. (1995). Electroencephalography and single photon emission computed tomography in dementia: a comparative study. *Psychol. Med.* 25, 631–638. doi: 10.1017/S0033291700033535
- Sporns, O. (2015). Cerebral cartography and connectomics. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 370:1668. doi: 10.1098/rstb.2014.0173
- Spreen, F., and Benton, A. (1977). *Manual of Instructions for the Neurosensory Center Comprehensive Examination for Aphasia*. Victoria, BC: University of Victoria.
- Staffen, W., Bergmann, J., Schönauer, U., Zauner, H., Kronbichler, M., Golaszewski, S., et al. (2009). Cerebral perfusion (HMPAO-SPECT) in patients with depression with cognitive impairment versus those with mild cognitive impairment and dementia of Alzheimer's type: a semiquantitative and automated evaluation. *Eur. J. Nucl. Med. Mol. Imaging* 36, 801–810. doi: 10.1007/s00259-008-1028-2
- Staffen, W., Schönauer, U., Zauner, H., Spindler, I., Maier, A., Iglseder, B., et al. (2006). Brain perfusion SPECT in patients with mild cognitive impairment and Alzheimer's disease: comparison of a semiquantitative and a visual evaluation. *J. Neural Transm.* 113, 195–203. doi: 10.1007/s00702-005-0321-5
- Stam, C. (2005). Nonlinear dynamical analysis of EEG and MEG: review of an emerging field. *Clin. Neurophys.* 116, 2266–2301. doi: 10.1016/j.clinph.2005.06.011
- Stögmann, E., Moser, D., Klug, S., Gleiss, A., Auff, E., Dal-Bianco, P., et al. (2016a). Activities of daily living and depressive symptoms in patients with subjective cognitive decline, mild cognitive impairment, and Alzheimer's disease. *J. Alzheimers Dis.* 49, 1043–1050. doi: 10.3233/JAD-150785
- Stögmann, E., Moser, D., Klug, S., Gleiss, A., Auff, E., Dal-Bianco, P., et al. (2016b). Impact of depressive symptoms on conversion from mild cognitive impairment subtypes to Alzheimer's disease: a community-based longitudinal study. *J. Alzheimers Dis.* 51, 405–415. doi: 10.3233/JAD-150603
- Stoppe, G., Staedt, J., Kögler, A., Schütze, R., Kunert, H., Sandrock, D., et al. (1995). 99mTc-HMPAO-SPECT in the diagnosis of senile dementia of Alzheimer's type—a study under clinical routine conditions. *J. Neural Transm. Gen. Sect.* 99, 195–211. doi: 10.1007/BF01271479
- Swan, A., Waddell, B., Holloway, G., Bak, T., Colville, S., Khan, Z., et al. (2015). The diagnostic utility of 99mTc-HMPAO SPECT imaging: a retrosPECTive case series from a tertiary referral early-onset cognitive disorders clinic. *Dement. Geriatr. Cogn. Disord.* 39, 186–193. doi: 10.1159/000369551
- Talbot, P., Lloyd, J., Snowden, J., Neary, D., and Testa, H. (1998). A clinical role for 99mTc-HMPAO SPECT in the investigation of dementia? *J. Neurol Neurosurg. Psychiatry* 64, 306–313. doi: 10.1136/jnnp.64.3.306
- Teipel, S., Grothe, M., Lista, S., Toschi, N., Garaci, F., and Hampel, H. (2013). Relevance of magnetic resonance imaging for early detection and diagnosis of alzheimer disease. *Med. Clin. North Am.* 97, 399–424. doi: 10.1016/j.mcna.2012.12.013
- Teipel, S., Grothe, M., Zhou, J., Sepulcre, J., Dyrba, M., Sorg, C., et al. (2016). Measuring cortical connectivity in Alzheimer's disease as a brain neural network pathology: toward clinical applications. *J. Int. Neuropsychol. Soc.* 22, 138–163. doi: 10.1017/S1355617715000995
- Thalman, B., Monsch, A., Schneitter, M., Bernasconi, F., Aebi, C., Zdenka, C., et al. (2000). The CERAD neuropsychological assessment battery (CERAD-NAB) - a minimal dataset as a common tool for German-speaking Europe. *Neurobiol. Aging* 21:30. doi: 10.1016/S0197-4580(00)82810-9
- Tranfaglia, C., Palumbo, B., Siepi, D., Sinzinger, H., and Parnetti, L. (2009). Semi-quantitative analysis of perfusion of Brodmann areas in the differential diagnosis of cognitive impairment in Alzheimer's disease, fronto-temporal dementia and mild cognitive impairment. *Hell J. Nucl. Med.* 12, 110–114. Available online at: <http://www.nuclmed.gr/wp/wp-content/uploads/2017/03/10.pdf>
- Valotassiou, V., Papatrifiayfyllou, J., Sifakis, N., Tzavara, C., Tsougos, I., Psimadas, D., et al. (2015). Clinical evaluation of brain perfusion SPECT with brodmann areas mapping in early diagnosis of Alzheimer's disease. *J. Alzheimers Dis.* 47, 773–785. doi: 10.3233/JAD-150068
- Van der Mussele, S., Fransen, E., Struyfs, H., Luyckx, J., Marien, P., Saerens, J., et al. (2014). Depression in mild cognitive impairment is associated with progression to Alzheimer's disease: a longitudinal study. *J. Alzheimers Dis.* 42, 1239–1250. doi: 10.3233/JAD-140405
- Van Heertum, R., Tikofsky, R., and Ichise, M. (2009). *Functional Cerebral SPECT and PET Imaging, 4th Edn*. Philadelphia, PA: Lippincott Williams & Wilkins.
- Vecchio, F., Miraglia, F., Curcio, G., Altavilla, R., Scarsia, F., Giambattistelli, F., et al. (2015). Cortical brain connectivity evaluated by graph theory in dementia: a correlation study between functional and structural data. *J. Alzheimers Dis.* 45, 745–756. doi: 10.3233/JAD-142484
- Vecchio, F., Miraglia, F., Marra, C., Quaranta, D., Vita, M., Bramanti, P., et al. (2014). Human brain networks in cognitive decline: a graph theoretical analysis of cortical connectivity from EEG data. *J. Alzheimer's Dis.* 41, 113–127. doi: 10.3233/JAD-132087
- Vecchio, F., Miraglia, F., Quaranta, D., Granata, G., Romanello, R., Marra, C., et al. (2016). Cortical connectivity and memory performance in cognitive decline: a study via graph theory from EEG data. *Neuroscience* 316, 143–150. doi: 10.1016/j.neuroscience.2015.12.036
- Welsh, K., Butters, N., Mohs, R., Beekly, D., Edland, S., Fillenbaum, G., et al. (1994). Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part V. A normative study of the neuropsychological battery. *Neurology* 44, 609–614. doi: 10.1212/WNL.44.4.609
- Winblad, B., Palmer, K., Kivipelto, M., Jelic, V., Fratiglioni, L., Wahlund, L., et al. (2004). Mild cognitive impairment - beyond controversies, towards a consensus: report of the International Working Group on Mild Cognitive Impairment. *J. Intern. Med.* 256, 240–246. doi: 10.1111/j.1365-2796.2004.01380.x
- World Health Organization (2016). *Dementia*. Fact Sheet Available online at: <http://www.who.int/mediacentre/factsheets/fs362/en/> (Accessed February 28, 2017).
- Wurtman, R. (2015). Biomarkers in the diagnosis and management of Alzheimer's disease. *Metabolism* 64, S47–S50. doi: 10.1016/j.metabol.2014.10.034
- Yeo, J., Lim, X., Khan, Z., and Pal, S. (2013). Systematic review of the diagnostic utility of SPECT imaging in dementia. *Eur. Arch. Psychiatry Clin. Neurosci.* 263, 539–552. doi: 10.1007/s00406-013-0426-z
- Zigmond, A., and Snaith, R. (1983). The hospital anxiety and depression scale. *Acta Psychiatr. Scand.* 67, 361–370. doi: 10.1111/j.1600-0447.1983.tb09716.x

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.

The reviewer MD and handling Editor declared their shared affiliation.

Copyright © 2017 Höller, Bathke, Uhl, Strobl, Lang, Bergmann, Nardone, Rossini, Zauner, Kirschner, Jahanbekam, Trinka and Staffen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Visualizing Hyperactivation in Neurodegeneration Based on Prefrontal Oxygenation: A Comparative Study of Mild Alzheimer's Disease, Mild Cognitive Impairment, and Healthy Controls

Kah Hui Yap^{1†}, Wei Chun Ung^{2†}, Esther G. M. Ebenezer¹, Nadira Nordin², Pui See Chin¹, Sandheep Sugathan³, Sook Ching Chan³, Hung Loong Yip³, Masashi Kiguchi⁴ and Tong Boon Tang^{2*}

¹ Medicine Based Department, Royal College of Medicine Perak, Universiti Kuala Lumpur, Kuala Lumpur, Malaysia, ² Centre for Intelligent Signal and Imaging Research, Universiti Teknologi Petronas, Seri Iskandar, Malaysia, ³ Community Based Department, Royal College of Medicine Perak, Universiti Kuala Lumpur, Kuala Lumpur, Malaysia, ⁴ Research & Development Group, Hitachi Ltd., Tokyo, Japan

OPEN ACCESS

Edited by:

Javier Ramírez,
University of Granada, Spain

Reviewed by:

Yvonne Höller,
Paracelsus Private Medical University
of Salzburg, Austria
Andres Ortiz,
University of Málaga, Spain

*Correspondence:

Tong Boon Tang
tongboon.tang@utp.edu.my

[†]These authors have contributed
equally to this work.

Received: 17 May 2017

Accepted: 17 August 2017

Published: 01 September 2017

Citation:

Yap KH, Ung WC, Ebenezer EGM, Nordin N, Chin PS, Sugathan S, Chan SC, Yip HL, Kiguchi M and Tang TB (2017) Visualizing Hyperactivation in Neurodegeneration Based on Prefrontal Oxygenation: A Comparative Study of Mild Alzheimer's Disease, Mild Cognitive Impairment, and Healthy Controls. *Front. Aging Neurosci.* 9:287. doi: 10.3389/fnagi.2017.00287

Background: Cognitive performance is relatively well preserved during early cognitive impairment owing to compensatory mechanisms.

Methods: We explored functional near-infrared spectroscopy (fNIRS) alongside a semantic verbal fluency task (SVFT) to investigate any compensation exhibited by the prefrontal cortex (PFC) in Mild Cognitive Impairment (MCI) and mild Alzheimer's disease (AD). In addition, a group of healthy controls (HC) was studied. A total of 61 volunteers (31 HC, 12 patients with MCI and 18 patients with mild AD) took part in the present study.

Results: Although not statistically significant, MCI exhibited a greater mean activation of both the right and left PFC, followed by HC and mild AD. Analysis showed that in the left PFC, the time taken for HC to achieve the activation level was shorter than MCI and mild AD ($p = 0.0047$ and 0.0498 , respectively); in the right PFC, mild AD took a longer time to achieve the activation level than HC and MCI ($p = 0.0469$ and 0.0335 , respectively); in the right PFC, HC, and MCI demonstrated a steeper slope compared to mild AD ($p = 0.0432$ and 0.0107 , respectively). The results were, however, not significant when corrected by the Bonferroni-Holm method. There was also found to be a moderately positive correlation ($R = 0.5886$) between the oxygenation levels in the left PFC and a clinical measure [Mini-Mental State Examination (MMSE) score] in MCI subjects uniquely.

Discussion: The hyperactivation in MCI coupled with a better SVFT performance may suggest neural compensation, although it is not known to what degree hyperactivation manifests as a potential indicator of compensatory mechanisms. However, hypoactivation plus a poorer SVFT performance in mild AD might indicate an inability to compensate due to the degree of structural impairment.

Conclusion: Consistent with the scaffolding theory of aging and cognition, the task-elicited hyperactivation in MCI might reflect the presence of compensatory mechanisms and hypoactivation in mild AD could reflect an inability to compensate. Future studies will investigate the fNIRS parameters with a larger sample size, and their validity as prognostic biomarkers of neurodegeneration.

Keywords: mild Alzheimer's disease, mild cognitive impairment, functional near-infrared spectroscopy, semantic verbal fluency task, prefrontal hemoglobin oxygenation

INTRODUCTION

The multiple dementia subtypes are associated with unique symptom patterns and brain abnormalities. Alzheimer's disease (AD) is a chronic progressive neurodegenerative brain disease that can occur in middle or old age. It represents the most common cause of dementia, accounting for 60–80% of cases (Alzheimer's Association, 2016). It causes increasing impairment in a range of cognitive functions that include memory, mood, reasoning, language, self-management, and behavior (Karantzoulis and Galvin, 2014). Mild cognitive impairment (MCI) is an intermediate state of clinical impairment, where the individuals affected have cognitive symptoms of a mild nature that are disproportionate to their age and education, while not meeting the criteria for dementia or AD (Petersen, 2009). Patients with MCI tend to progress toward developing AD at a rate of ~15% per year (Gauthier et al., 2006).

Magnetic resonance imaging (MRI) studies have demonstrated that deficits found in patients with AD are associated with volumetric changes in the prefrontal cortex (PFC) (Salat et al., 2001; McNab and Klingberg, 2008; Zanto et al., 2011). It has been suggested that the cognitive decline related to normal aging is attributable to a reduction in white matter rather than gray matter (Marner et al., 2003). In contrast, significant loss of gray matter, which is composed of cortical neurons and glia has been seen in AD, and leads to a reduced neuronal activation in the PFC of AD when compared with HC (Yankner et al., 2008). MCI is positioned between mild AD and normal cognitive aging (Sperling, 2007). While age-related regional volume loss is apparent and widespread in normal cognitive aging, a unique pattern of structural vulnerability, reflected in differential volume loss in specific regions, has been identified in patients with MCI (Driscoll et al., 2009). Also, patients with AD perform more poorly in semantic tasks associated with compromised activation in the left PFC when assessed using functional MRI (fMRI) (Johnson et al., 2000). The

hippocampus is among the first non-cortical regions affected by AD-related neurodegeneration; hippocampal atrophy has been associated with early memory decline in MCI and AD (Barnes et al., 2007; Shi et al., 2009; Nho et al., 2012). Several previous studies of MCI patients and healthy older adults reported reduced activation in the hippocampus and PFC of the former (Johnson et al., 2006; Dannhauser et al., 2008; Mandzia et al., 2009). These results suggest that AD may be characterized by reduced brain activation due to the pathological changes associated with it.

One key element influencing these early changes is neural compensation. The capacity for neural compensation is inversely proportional to the severity of neurodegeneration (Price and Friston, 2002). As neural damage worsens, both cognitive processing efficiency and capacity become impeded. This affects neuronal recruitment capacity, reduces compensation capacities, and ultimately results in poorer performance (Scarmeas et al., 2003). As an example, transcranial magnetic stimulation has been shown to induce compensatory activation associated with information retrieval when applied to either the right or left PFC of patients with severe cognitive impairment. However, similar activation was only observed when applied to the left PFC of healthy individuals. These data suggest that right PFC recruitment acts as one of the functional compensatory mechanisms in cognitively impaired individuals (Cotelli et al., 2008). In addition, neural compensation prolonged the period of MCI and delayed progression to AD (Baazaoui et al., 2017), implying that such compensatory mechanisms may play a more crucial role in MCI than in AD. In the present study, the aim was to investigate the difference between MCI and AD using functional near-infrared spectroscopy (fNIRS) to identify hyperactivation as a potential indicator of the occurrence of compensation.

fNIRS is an emerging technology in neuroimaging that has been increasingly employed in the past 20 years (Ferrari and Quaresima, 2012). It uses an optical window where the head scalp is almost transparent to near infrared light at wavelengths of 700–900 nm, and the fact that oxygenated- and deoxygenated-hemoglobin (oxy-Hb and deoxy-Hb) are strong absorbers of light (Heinzel et al., 2013). Based on the concept of neurovascular coupling, neuronal activities are measured by fNIRS as changes in oxy-Hb and deoxy-Hb concentrations using the Modified Beer-Lambert Law. fNIRS has, so far, been applied to many psychiatric disorders, for example, to differentiate depression, bipolar disorder, and schizophrenia (Fallgatter et al., 1997; Arai et al., 2006; Takizawa et al., 2014). The technique has a good

Abbreviations: AD, Alzheimer's disease; aMCI, amnesic mild cognitive impairment; A β , amyloid β -protein; CDR, clinical dementia rating; deoxy-Hb, deoxygenated-hemoglobin; fMRI, functional magnetic resonance imaging; fNIRS, functional near-infrared spectroscopy; HC, healthy controls; MCI, Mild cognitive impairment; MMSE, mini-mental state examination; MRI, magnetic resonance imaging; naMCI, Non-amnesic mild cognitive impairment; oxy-Hb, oxygenated-hemoglobin; PFC, prefrontal cortex; SVFT, semantic verbal fluency task; $N_{\text{left PFC/right PFC}}$: the number of activated measurement channels; $\Delta \text{oxy-Hb}_{\text{left PFC/right PFC}}$: the change in oxy-Hb concentration during the activation period; $t_{\text{left PFC/right PFC}}$: the time taken to achieve activation level; $m_{\text{left PFC/right PFC}}$: the slope in the first 5 s after task onset.

temporal resolution (~ 1 ms) and a reasonable spatial resolution (~ 1 cm). In addition, it has several advantages over the more widely-used fMRI, including lower costs, portability, and low levels of subject constraint (Ehlis et al., 2014). It has been demonstrated that fNIRS readings are highly correlated with fMRI with respect to measuring cognitive tasks (Cui et al., 2011). More specifically, multiple fNIRS-fMRI studies have consistently shown increases in oxy-Hb in the left PFC during semantic verbal fluency tasks (SVFT) (Fallgatter et al., 1997; Heinzel et al., 2013; Wagner et al., 2014; Gutierrez-Sigut et al., 2015). So, fNIRS represents a potential alternative technique to fMRI for the investigation of the differences in neuronal activities between individuals with differing cognitive impairments.

fNIRS has been used in previous studies into AD and MCI. Patients with AD were identified to have lower activation of the PFC compared to healthy controls (HC) during both letter and SVFT, while HC performed better in the tasks. Differences in hemodynamic responses were predominantly found in the left hemisphere, supporting the idea that good performances in verbal fluency tasks are associated with higher left hemisphere activation, predominantly (Fallgatter et al., 1997; Arai et al., 2006; Herrmann et al., 2008). This poorer performance in patients with AD has been attributed to the loss of hemispheric asymmetry rather than to the level of PFC activation alone (Fallgatter et al., 1997). The effect of hemispheric asymmetry or lateralization plays a more crucial role in MCI. The rightward shift of frontal activations in the MCI group might reflect the presence of cortical reorganization; the recruitment of the right PFC has been suggested to compensate the loss in the left PFC (Yeung et al., 2016). Another study has shown that compared to HC, patients with AD demonstrated lower activation of the frontal and bilateral parietal areas, while patients with MCI have lower right parietal activation (Arai et al., 2006). These results suggest that more detailed studies into neural compensatory mechanisms are warranted.

In the present study, fNIRS was used with a wide coverage of the PFC to investigate hyper/ hypoactivation in MCI, mild AD, and HC. The PFC was selected as the region of interest as the PFC is not only involved in semantic memory, but is also accessible by fNIRS, which can penetrate brain tissue up to a depth of 5 cm in the cortical region (Ranger et al., 2011). SVFT was selected as the cognitive activation task based on previous studies of semantic memory (Fallgatter et al., 1997; Thompson-Schill et al., 1997; Perry et al., 2000; Grossman et al., 2002; Heinzel et al., 2013; Wagner et al., 2014; Gutierrez-Sigut et al., 2015; Yeung et al., 2016). In accordance with the compensation theory (Yankner et al., 2008), it was expected that neural compensation would be observed in MCI, but not in AD during the tasks. Patients with MCI suffer from a relatively small degree of neurodegeneration compared to patients with AD, therefore, it was predicted that MCI patients would be able to activate neural compensation, as manifested by brain hyperactivation, in order to maintain task performance. As patients with AD suffer from more severe neuronal damage, to the extent that their neural compensation abilities might be compromised, it was thought that hypoactivation would be observed instead. We aimed to investigate to what degree

hyperactivation, as a potential indicator for compensation, manifests in MCI and to what degree it is compromised in AD. Our first hypothesis was that during SVFT, subjects with normal cognitive aging (the HC group) would perform better, followed by MCI and AD. Secondly, we expected hyperactivation and hypoactivation in the PFC of MCI and AD, respectively, when subjects were tested with a semantic memory task.

METHODS

Participants

We recruited participants, who were right-handed and able to converse in English, through a local dementia day-care center as well as from the local community where English was the common medium of instruction in the past. Patients with MCI and patients with mild AD were recruited through purposive sampling with group-specific inclusion criteria. This was followed by the recruitment of HC who were age (± 2 years), gender- and education-matched. We assessed the participants, ruling out anyone with a psychiatric disorder. Additionally, neurological disorders, including other forms of dementia were excluded. Assessments were performed by a psychiatrist using an evaluation of medical history, including careful examination of the course of progression, the relative salience of cognitive, behavioral and physical symptoms and signs, and patterns of cognitive impairment. Additionally, mental state was examined using a Mini-Mental State Examination (MMSE), which is a 30-point questionnaire providing a quantitative measure of cognitive status or cognitive impairment (Folstein et al., 1975). Other exclusion criteria were other medical diagnoses affecting cognitive functioning, including kidney failure, stroke, known lesions, and any history of significant trauma. The study protocol was approved by the Medical Research Ethics Committee of the University of Kuala Lumpur (Approval no.: 2015/032). All participants were briefed about the nature of the experimental procedures prior to providing demographic information and written informed consent in accordance with the Declaration of Helsinki. All tests and experiments were completed on the same day, with a short break between the test and the experiment. Both healthy participants and patients were remunerated for their participation.

Clinical Measures

We used the Clinical Dementia Rating (CDR), an observer rating scale designed to rate the severity of dementia (Morris, 1993), for the diagnosis of dementia and group allocation. CDR scores of 0, 0.5, and 1 were assigned to HC, MCI, and mild AD, respectively; participants with CDR scores of 2 (moderate) and 3 (severe) were excluded, as this study focused on MCI and mild AD. In addition, we assessed and assigned each participant an MMSE score. CDR has a moderate to high inter-rater reliability of 0.62 (Rockwood et al., 2000). MMSE has high inter-rater reliability, ranging between 0.82 and 0.91 (Magni et al., 1996). We matched the HC to those of the combined sample of patients according to age, sex, and education.

fNIRS Technology

Throughout this study, a multichannel OT-R40 fNIRS topography system (Hitachi Medical Corporation, Japan) was employed to measure the brain activity at a sampling rate of 10 Hz. Changes in oxy-Hb and deoxy-Hb signals were measured in units of mM·mm. fNIRS has been reported to be more sensitive to gray matter when a larger source-detector separation (up to ~4.5 cm) is used, albeit at the expense of both spatial resolution and partial pathlength factors (Strangman et al., 2013). Taking everything into consideration, the source-detector distance was fixed at 3 cm, within the suggested optimal range for adult heads (3–3.5 cm) (Li et al., 2011). The midpoint between pairs of sources and the detector was defined as a measurement channel. The probes were arranged into a 3×11 layout (see **Figure 1A**) to form a total of 52 measurement channels that were sufficient to measure the entire PFC and part of the temporal cortex (see **Figure 1B**; Ishii-Takahashi et al., 2014; Takizawa et al., 2014). According to the international 10–20 system (Klem et al., 1999), source no. 23 and 28 were positioned directly at T4 and T3, respectively. The probes were attached to a flexible head cap, which was relatively easy, fast and convenient to wear. All channels were checked to ensure that the probes were in contact with the scalp. The entire set-up process took an average of <10 min.

Task Paradigm

Participants were seated comfortably in a working chair and were instructed to avoid movement and to place their hands on the armrests during the experiment. Standardized verbal instructions and explanations regarding the tasks were given in English. Prior to any new measurement, practice was given, allowing the participants to familiarize themselves with the experimental procedures. SVFT was selected as the cognitive activation task in this study. During fNIRS measurements, participants were instructed to provide as many words verbally as possible from a particular category (Fruits, Food, and Animals). The experimental session was preceded by 20 s of pre-task rest period. Each category lasted 60 s and was followed by 20 s of rest. Participants were told to avoid repeating the same word and they were asked to keep their eyes on the LCD screen for the entire task period, which lasted for a total of 260 s (see **Figure 2**).

Data Analysis

fNIRS Data

Oxy-Hb was selected as the focus of measurements due to its sensitivity to task-associated changes (Sato et al., 2006; Cui, 2011). Probes that were not in good contact with the scalp may have resulted in rapid large changes (such as high amplitude spikes) in oxy-Hb signals. An fNIRS channel was considered to be “noisy” if there were very large spikes (changes in oxy-Hb

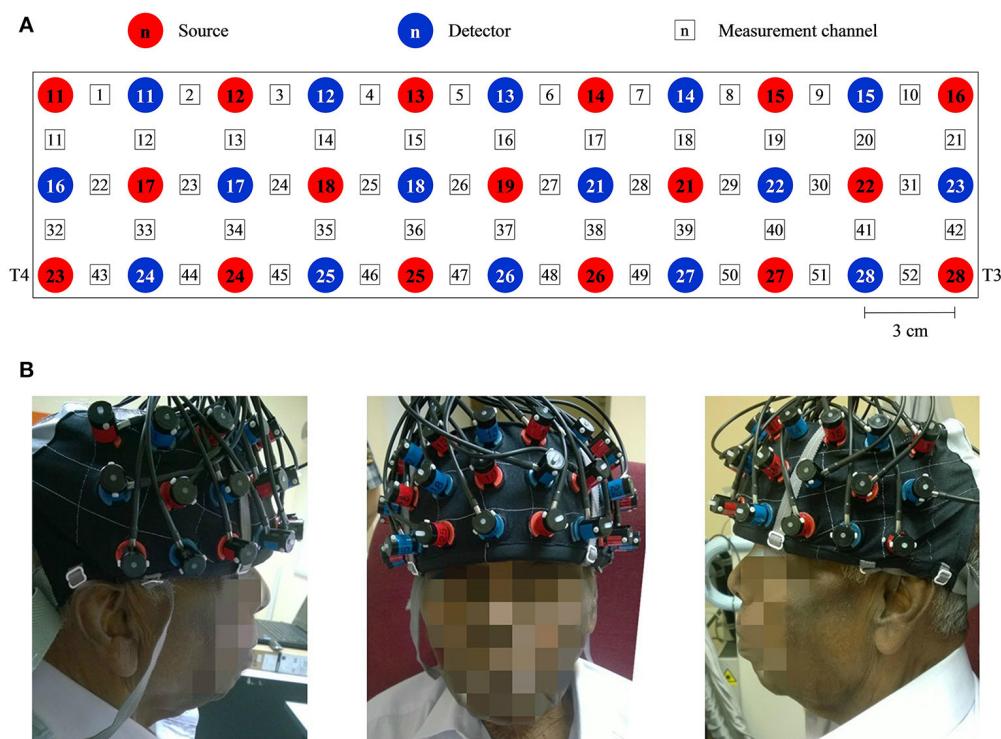


FIGURE 1 | A system consisting of 52 measurement channels was used. **(A)** The probes were arranged to 3×11 layout. The source-detector distance was fixed at 3 cm and the space between pairs of source and detector was defined as a measurement channel. **(B)** One of the participants wearing the flexible head cap which housed the probes. Source no. 23 and 28 were positioned directly at T4 and T3 accordingly to the international 10–20 system. Consent was obtained from the individual for the publication of this image.

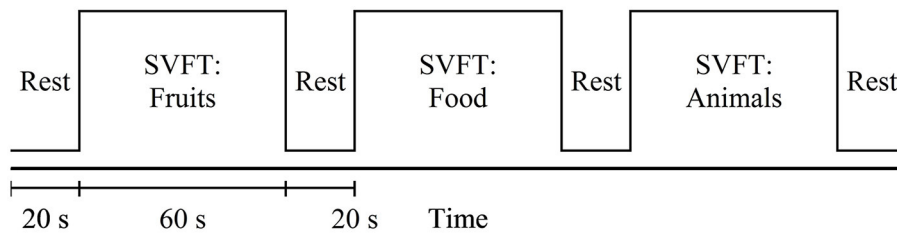


FIGURE 2 | Schematic diagram of the measurement sequence: SVFT = semantic verbal fluency test. It started with a 20 s rest. Each category of SVFT was 60 s long and was followed by 20 s of rest.

larger than 0.5 mM·mm in amplitude) and all noisy channels were excluded from subsequent analyses. All data analyses were performed using a software Platform for Optical Topography Analysis Tools (Sutoko et al., 2016). A Butterworth bandpass filter with cut-off frequencies of 0.01–0.8 Hz was applied to remove instrumental or physiological noise (Luu and Chau, 2009).

Three categories of SVFT were represented by three 100-s blocks of data for analysis. As discussed in the task paradigm, each block consisted of a 20-s pre-task rest period, a 60-s task period followed by a 20-s post-task period (see to **Figure 3**). A moving average filter with a window size of 50 data points (5 s) was applied to remove high frequency noise from the measured oxy-Hb signals (Ishii-Takahashi et al., 2014). With respect to the start of each block, the oxy-Hb signals were then baseline-corrected to a zero baseline. Blocks with a large spike noise in the oxy-Hb signal (changes in oxy-Hb larger than 0.5 mM·mm in amplitude) were excluded from further analyses. Subsequently, fNIRS signals from each channel were meaned over all the remaining blocks so that each subject should have had only one average fNIRS signal per channel. Subsequent analyses were performed using these average fNIRS signals. Previous fNIRS studies have suggested that earliest activation starts from 5 s after task onset (Sato et al., 2006) and sharp increases in activation are often observed at around 5–10 s after onset (Maki et al., 1995). Therefore, the level of activation at each channel was determined for each individual using the percentage signal change, which was calculated using the following formula:

$$\text{percent signal change} = \frac{\text{oxy} - \text{Hb}_{t=5:65}(\text{avg}) - \text{oxy} - \text{Hb}_{t=-10:0}(\text{avg})}{|\text{oxy} - \text{Hb}_{t=-10:0}(\text{avg})|} \times 100\% \quad (1)$$

where $\text{oxy} - \text{Hb}_{t=5:65}(\text{avg})$ is the average oxy-Hb signal during the task period, after accounting for hemodynamic delay, and $\text{oxy} - \text{Hb}_{t=-10:0}(\text{avg})$ is the average oxy-Hb signal during the rest period (–10–0 s of the pre-task rest period; see **Figure 3**). Channels located in the PFC that showed a percentage signal change of larger than 50% were empirically considered to be activated and were further divided into left and right PFC. Hence, it was possible for each subject to have a different number of

activated channels, but they were all within the regions of interest (the left and right PFC).

Due to hemodynamic delay, the period between 5 and 25 s after task onset was defined as the activation period (see **Figure 3**). The activation signal was defined as the difference between the average oxy-Hb signal during the activation period and during the rest period. For each participant, the mean activation signal in the left and right PFC was calculated by averaging the signals obtained in activated measurement channels located in the left and right PFC, respectively. Based on the mean activation signals, for both the left and right PFC, the time taken to achieve the activation level was determined and the slope in the first 5 s after task onset was calculated using the following formula:

$$\text{slope} = \frac{\text{oxy} - \text{Hb}_{\text{left PFC/right PFC}}^{t=5} - \text{oxy} - \text{Hb}_{\text{left PFC/right PFC}}^{t=0}}{5} \times \frac{\text{mM} \cdot \text{mm}}{\text{s}} \quad (2)$$

where $\text{oxy} - \text{Hb}_{\text{left PFC/right PFC}}^{t=0}$ and $\text{oxy} - \text{Hb}_{\text{left PFC/right PFC}}^{t=5}$ denote the mean activation signals in the left or right PFC at the onset of the task and 5 s after the onset of the task, respectively.

The number of activated measurement channels, the mean activation signals in both the left and right PFC, and the time taken to achieve the activation level and the slope in the first 5 s after task onset were then statistically assessed.

Statistics

Differences between all three groups in MMSE scores were first tested using the multiple non-parametric Mann–Whitney U tests without multiple testing correction, followed by Mann–Whitney U tests with Bonferroni–Holm correction. Both tests were repeated to assess whether there were any within-category group differences in the number of words generated in the three categories of SVFT. The Bonferroni–Holm correction would test each individual hypothesis in a sequential rejective manner at $\alpha / [n - \text{rank number of the pair (by degree of significance)} + 1]$, where α is the desired significance level (0.05) and n is the number of comparisons. fNIRS data that were statistically assessed included the number of activated measurement channels, the mean activation signals in both the left and right PFC, the time taken to achieve the activation level and the slope in the first 5 s after task onset. For each of these four

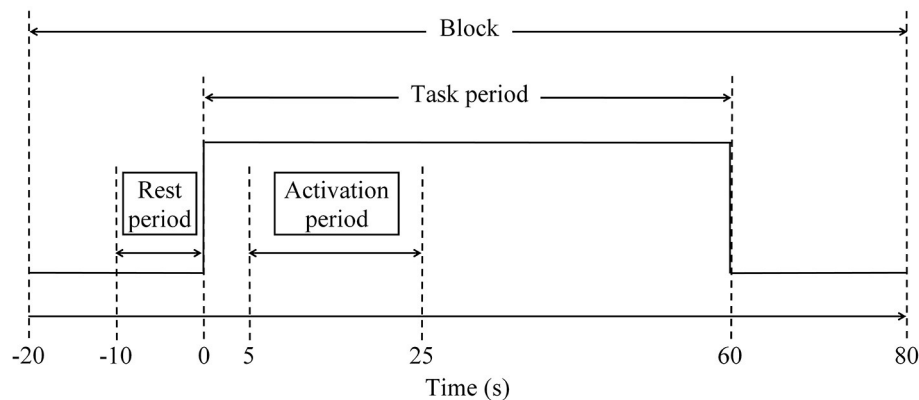


FIGURE 3 | Schematic diagram of a block to facilitate analysis. Each block consisted of 20-s pre-task rest period, 60-s task period and 20-s post-task period. Three categories of SVFT can be represented as three 100-s blocks.

fNIRS parameters, statistical significance was estimated using multiple comparisons between the three different groups at each category level, with two sets of independent two-sample *t*-tests—one without multiple testing correction and one with similar Bonferroni-Holm correction. Finally, to investigate the relationship between the MMSE score and various fNIRS parameters, correlation and simple linear regression analyses were performed using the MMSE score as a continuous independent variable.

RESULTS

Sample Characteristics

We excluded the data collected from one patient with moderate AD (CDR Score = 2) as well as two left-handed participants (one each from HC and MCI, as we were focusing only on right-handed participants). The final group of participants consisted of 31 HC and 30 patients (MCI: 12; mild AD: 18) matched for age, sex, education, and handedness. Demographic information about age, gender and education level was collected (see **Table 1**).

Behavioral Data

The behavioral results (MMSE scores and the number of words generated in the three categories of SVFT) are summarized in **Table 1** and illustrated in **Figure 4**. The number of comparisons for MMSE scores is three (for the three groups) and for SVFT is nine (3 groups \times 3 categories). There was a significant difference in MMSE score between the groups (see **Table 1** and **Figure 4A**; *p*: HC vs. MCI = 0.0033, HC vs. mild AD < 0.0001, MCI vs. mild AD = 0.0012). HC had the highest MMSE scores, followed by MCI, then mild AD. Referring to **Figure 4B**, significant group differences in the number of words given in the “Fruits” category (*p*: HC vs. MCI = 0.0081, HC vs. mild AD < 0.0001, MCI vs. mild AD = 0.0084) and the “Food” category (*p*: HC vs. MCI = 0.0094, HC vs. mild AD < 0.0001, MCI vs. mild AD = 0.0079) were found. For the “Animals” category, the number of words given by HC was significantly higher than for mild AD (*p* < 0.0001). In comparison to MCI, HC produced more words, but it was not

statistically significant (*p* = 0.0217). The number of words given by MCI was also higher than mild AD, but the difference was not statistically significant (*p* = 0.1669).

fNIRS Data

To characterize the fNIRS responses, we derived the following parameters, as shown in **Figure 5**:

- (i) the number of activated measurement channels, $N_{\text{left PFC/right PFC}}$
- (ii) the change in oxy-Hb concentration during the activation period, $\Delta\text{oxy} - \text{Hb}_{\text{left PFC/right PFC}}$
- (iii) the time taken to achieve the activation level, $t_{\text{left PFC/right PFC}}$
- (iv) the slope in the first 5 s after task onset, $m_{\text{left PFC/right PFC}}$

As illustrated in **Figure 5A**, $N_{\text{left PFC}}$ was found to be higher (but not significantly so) than $N_{\text{right PFC}}$ in both HC and MCI. However, mild AD showed a higher $N_{\text{right PFC}}$, compared to the $N_{\text{left PFC}}$. The oxy-Hb signals measured in these activated measurement channels were then averaged across each group of participants to obtain an overall signal for the left and right PFC (see **Figure 6**).

Figure 5B shows the $\Delta\text{oxy} - \text{Hb}_{\text{left PFC/right PFC}}$ for all groups. The highest $\Delta\text{oxy} - \text{Hb}_{\text{left PFC}}$ was observed in MCI followed by HC, while mild AD showed the lowest $\Delta\text{oxy} - \text{Hb}_{\text{left PFC}}$. A similar trend was also observed in the right PFC. However, in both the left and right PFC there were no significant differences between all three groups with respect to the level of activation. Nevertheless, all three groups demonstrated a similar trend, with higher $\Delta\text{oxy} - \text{Hb}_{\text{right PFC}}$ compared to $\Delta\text{oxy} - \text{Hb}_{\text{left PFC}}$, although differences were not statistically significant.

The $t_{\text{left PFC/right PFC}}$ was calculated and is illustrated in **Figure 5C**. When no multiple testing correction was used, the $t_{\text{left PFC}}$ by HC was shorter than MCI (*p* = 0.0047) and mild AD (*p* = 0.0498). On the other hand, mild AD took a longer time than HC (*p* = 0.0469) and MCI (*p* = 0.0335) in the right PFC. The $m_{\text{left PFC/right PFC}}$ was also determined (see **Figure 5D**). Both HC and MCI showed steeper $m_{\text{right PFC}}$ when compared to mild

TABLE 1 | Participants' demographic information and the pairwise Mann-Whitney U-test results.

| Characteristic | HC (n = 31) | MCI (n = 12) | mild AD (n = 18) | Mann-Whitney U-test | | |
|------------------------|-------------|--------------|------------------|---------------------|--------------------|-------------------|
| | Mean (SD) | Mean (SD) | Mean (SD) | p-value (r) | | |
| | | | | HC vs. MCI | HC vs. mild AD | MCI vs. mild AD |
| Age, years | 72.6 (8.5) | 73.1 (8.2) | 74.7 (10.0) | | | |
| Gender, M/F | 19/12 | 8/4 | 12/6 | | | |
| Education level, P/S/T | 3/18/10 | 1/7/4 | 3/12/3 | | | |
| CDR rating | 0 | 0.5 | 1 | | | |
| MMSE score | 28.7 (1.5) | 26.0 (3.1) | 21.2 (3.6) | 0.0033** (0.4485) | <0.0001** (0.7957) | 0.0012** (0.5896) |
| SVFT, words | | | | | | |
| Fruits | 13.6 (4.9) | 9.0 (3.7) | 5.7 (2.2) | 0.0081** (0.4036) | <0.0001** (0.6934) | 0.0084** (0.4809) |
| Food | 14.6 (4.9) | 10.2 (3.9) | 6.3 (2.7) | 0.0094** (0.3961) | <0.0001** (0.7459) | 0.0079** (0.4851) |
| Animals | 15.2 (3.7) | 10.8 (5.7) | 7.8 (3.0) | 0.0217** (0.3500) | <0.0001** (0.7304) | 0.1669 (0.2524) |

** $p < 0.05$ with Bonferroni-Holm correction.

HC, healthy control; MCI, mild cognitive impairment; mild AD, mild Alzheimer's disease; r, effect size; SD, standard deviation; M, male; F, female; P, primary; S, secondary; T, tertiary.

AD ($p = 0.0432$ and 0.0107 , respectively). In summary, HC used a shorter $t_{\text{left PFC}}$ than MCI and mild AD, and mild AD took longer $t_{\text{right PFC}}$ than HC and MCI, while HC and MCI demonstrated a steeper $m_{\text{right PFC}}$ than mild AD. However, the results described above were not significant when a Bonferroni-Holm correction was applied (number of comparisons is $3 \text{ groups} \times 4 \text{ fNIRS parameters} = 12$).

We also investigated, using simple linear regression, if there was any correlation between fNIRS parameters, i.e., $N_{\text{left PFC/right PFC}}$, $t_{\text{left PFC/right PFC}}$ and $m_{\text{left PFC/right PFC}}$, with behavioral data (MMSE scores). Interestingly, there was a moderate positive correlation ($R \geq 0.5$) between one of the fNIRS parameters and the MMSE score in MCI, but not in the case of HC or mild AD. More specifically, in the left PFC, $\Delta\text{oxy} - Hb_{\text{left PFC}}$ was moderately correlated to the MMSE score ($R = 0.5886$), as illustrated in **Figure 7**; **Table 2** summarizes the results.

Finally, we noticed a large degree of inter-subject variation in fNIRS parameters, as shown in **Figure 5**. Such high variance not only makes the small group differences hard to distinguish statistically, but also results in a large overlap between the data from all three groups.

DISCUSSION

Behavioral Data

This study was designed to investigate the differences in prefrontal oxygenation between normal cognitive aging, MCI and mild AD using fNIRS. On a behavioral level, HC performed better than MCI, followed by mild AD in all three categories of SVFT, except for with the "Animals" category between HC and MCI; here, no significant difference was found between MCI and mild AD. We also observed that participants, regardless of study group, were relatively faster in providing responses during the "Animals" category compared to the "Food" and "Fruits" categories. This finding might be due to all three groups being more familiar with the names of animals. This has been

reported in previous studies, which have shown that naming is influenced by item frequency and familiarity (Patterson and Hodges, 1992; Lambon Ralph et al., 1998). It is possible that this difference between categories may be related to the varying degree of distinctive features among category members (Moss et al., 2002). This will not be elaborated upon further here, as it is not the main focus of the present study. Overall, these results were consistent with past studies that have utilized SVFT in various neuroimaging modalities (Fallgatter et al., 1997; Heinzel et al., 2013; Wagner et al., 2014; Yeung et al., 2016). This suggests that the SVFT is a reliable cortical activation task to be used in conjunction with fNIRS measurements. Here, it was observed that patients diagnosed with a greater degree of dementia, i.e., with higher CDR scores, gave repeated words more frequently. They had a tendency to forget which words they had already given, an attribute of a deteriorated right PFC where the working memory is for monitoring, or keeping immediate information on-line during tasks (Hayama and Rugg, 2009). The number of repetitions in SVFT and such attributes will not be discussed further here.

fNIRS Data

$t_{\text{left PFC/right PFC}}$ and $m_{\text{left PFC/right PFC}}$

The results suggested that there were two fNIRS parameters that were significantly different between groups when the subjects were engaged in SVFT, only when no multiple testing correction was used. The first parameter is the $t_{\text{left PFC/right PFC}}$. In the left PFC, HC took a shorter time in achieving the targeted activation level, compared to patients with MCI and mild AD. However, contrasting with the belief that MCI would have a faster activation than mild AD, patients with mild AD actually took a shorter time. Conversely, with respect to the right PFC, MCI demonstrated the shortest time taken, followed by HC and mild AD. This is suggestive of the faster hemodynamic response in the right PFC of the MCI possibly being a compensatory response for the loss of the left PFC (Yeung et al., 2016). Taken separately, the poorer performance in SVFT and smaller

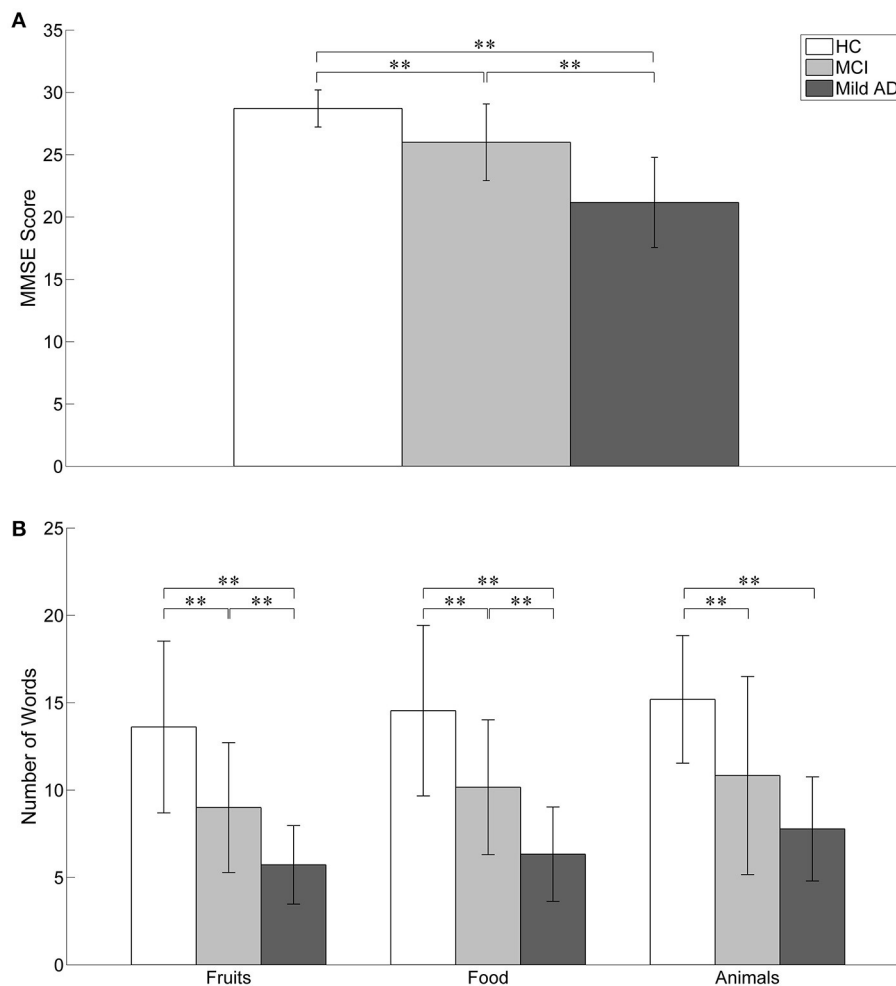


FIGURE 4 | The behavioral results. Statistical analysis was performed used two-sample *t*-test. $**p < 0.05$ with Bonferroni-Holm correction. The error bars represent the standard deviations. **(A)** The MMSE score of HC was significantly higher than MCI and mild AD. MCI also achieved significant higher MMSE score as compared to mild AD. **(B)** In the SVFT, the performance was measured by the number of words given. Significant group differences between all three groups were found in the number of words given in “Fruits” and “Food” categories. For “Animals” category, the number of words given by HC was significantly higher than mild AD. In comparison to MCI, HC produced more words but it was not statistically significant. The number of words given by MCI was also higher than mild AD but the difference was not significant.

activation in conjunction with the shorter time taken for left PFC activation in patients with mild AD might suggest that the compensatory mechanism is compromised. The second parameter examined was the $m_{\text{left PFC/right PFC}}$. Experimental results showed that the $m_{\text{right PFC}}$ was significantly greater in MCI compared to mild AD, further suggesting its importance to describing compensatory mechanisms. While the Bonferroni-Holm correction demonstrated that there were no significant differences, it is worthwhile explaining here the underlying mechanisms for these results.

Two possible explanations have been proposed for the underlying mechanism. The first explanation was consistent with that of the scaffolding theory of aging and cognition, in which additional circuitry is recruited to support declining brain function that has become inefficient (Park and Bischof, 2013).

This is commonly manifested in older adults that show increased contralateral right PFC recruitment for both working memory and episodic encoding (Reuter-Lorenz et al., 2000; Cabeza et al., 2002), which is consistent with our results. Such bilateral activation may be a form of interhemispheric interaction that has been claimed to be vital in neural compensatory mechanisms (Banich, 1998; Cabeza, 2002). Additionally, the results presented here imply that compensation and neuroplasticity might be present in the PFC of MCI, but not in mild AD. It has also been suggested that such compensatory ability might be reduced or lost in the progression of MCI toward AD, as neurodegeneration suffered by AD patients is severe enough to halt natural compensation (Clement et al., 2013). This is supportive of the results presented here, where patients with mild AD tended to forget which words they had already given added to

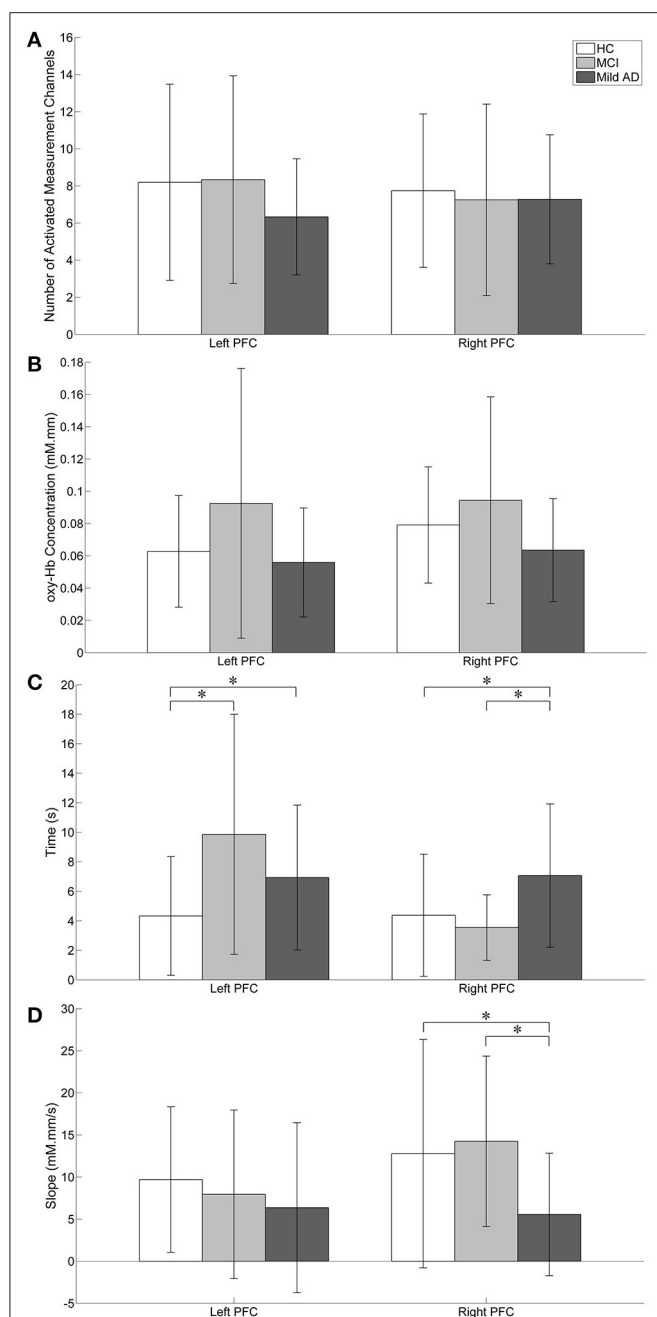


FIGURE 5 | Visual representation of fNIRS data. Statistical analysis was performed using two-sample *t*-test. **p* < 0.05 without multiple testing correction. The error bars represent the standard deviations. **(A)** The number of activated measurement channels: $N_{\text{left PFC/right PFC}}$; there were no significant differences between groups in both the left and right PFC. **(B)** oxy-Hb concentration change during the activation period: $\Delta\text{oxy} - \text{Hb}_{\text{left PFC/right PFC}}$; Higher activation was observed in MCI followed by HC while mild AD showed the least in both the left and right PFC. The right PFC was more activated than the left PFC in all groups. For both the left and right PFC, there were no significant group differences. **(C)** Time taken to achieve activation level: $t_{\text{left PFC/right PFC}}$; HC's $t_{\text{left PFC}}$ was significantly shorter than MCI. **(D)** Slope in the first 5 s after task onset: $m_{\text{left PFC/right PFC}}$; in comparison to mild AD, MCI showed significantly steeper $m_{\text{right PFC}}$.

the lack of activation of their right PFC, which is responsible for monitoring the immediate information during a task (Hayama and Rugg, 2009). This is further supported as the right PFC is specifically involved in semantic aspects of lexico-semantic processing (Joanette and Goulet, 1986).

We propose further, that the additional right PFC involvement and poorer memory performance can be explained using the inhibitory hypothesis. According to the inhibitory hypothesis, the non-dominant right PFC is normally suppressed by its dominant contra-lateral counterpart (Cox et al., 2015). Such transcallosal inhibition might be impaired with left PFC or anterior corpus callosum atrophy, which might result in extra recruitment of the right PFC. It has been found that additional non-dominant right PFC activity might reflect age-related changes in the brain and has been reported to be negatively correlated with memory performance (de Chastelaine et al., 2011). However, it has also been suggested that such disinhibition could reflect an attempted compensatory process, which is insufficient to fully compensate for age-related neurodegeneration. Our results point more toward the first explanation as we found that MCI, with right PFC activation relatively higher than both HC and mild AD, actually performed better than the latter in the SVFT. Hence, we suggest that the identification of a compensatory role for the right PFC might offer a potential target area for neurorehabilitation (Cotelli et al., 2008). It is expected that in the future the residual plasticity in the right PFC of cognitively impaired patients, particularly those with MCI, might be effectively harnessed by neurorehabilitation and other interventional techniques. At this juncture, it is necessary to examine other parameters to identify differences in task-related activities across different populations.

$N_{\text{left PFC/right PFC}}$

Considering the number of activated fNIRS channels, there were no significant differences between both the right and left PFC for all three groups. The left PFC, which is responsible for semantic memory, was activated in the SVFT, as expected (Grossman et al., 2002). The activation of the right PFC, however, may be suggestive of participants being engaged in object imagery prior to recalling names during the task. In addition, it might also indicate ongoing monitoring of semantic information (Hayama and Rugg, 2009); the episodic memory located in the right PFC was engaged to ensure that participants did not repeat names during the task. As compared with HC, a smaller region of right PFC activation was found in the other two groups (MCI and mild AD). This could indicate the lack of a monitoring process, thus partially explaining poorer performance in SVFT in these groups compared to HC. However, this cannot explain the differences in SVFT performances between MCI and mild AD.

$\Delta\text{oxy} - \text{Hb}_{\text{left PFC/right PFC}}$

MCI demonstrated the greatest oxygenation levels of PFC activation, in both the left and right hemispheres, during SVFT, followed by HC, and mild AD. Despite not being statistically significant, this result might be of clinical significance. The result is in agreement with previous studies, utilizing various neuroimaging modalities, which have demonstrated similar

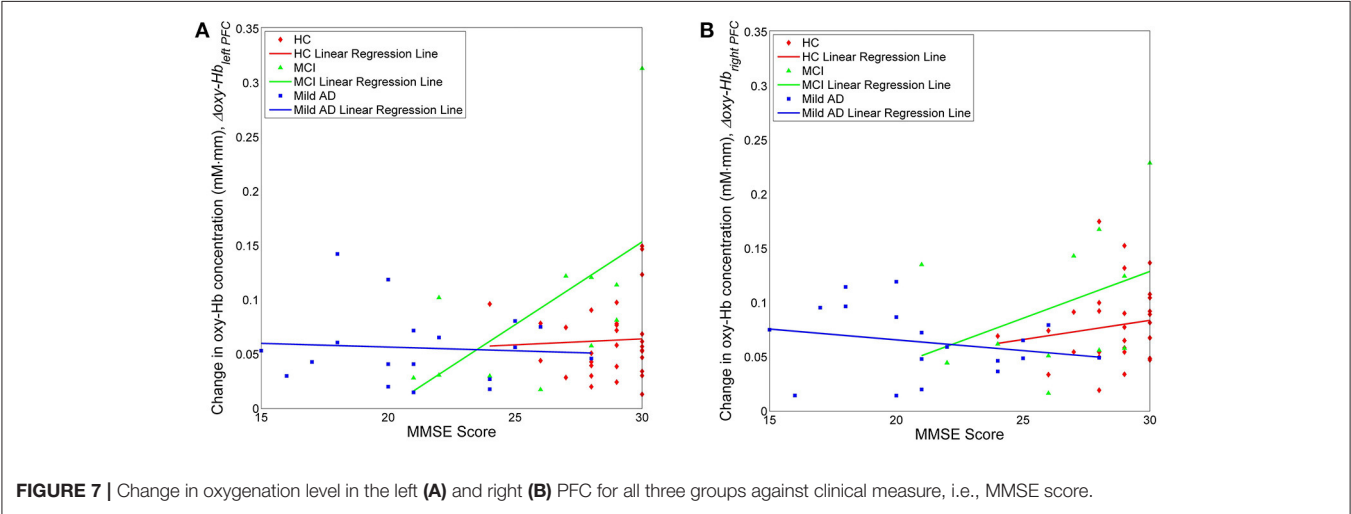
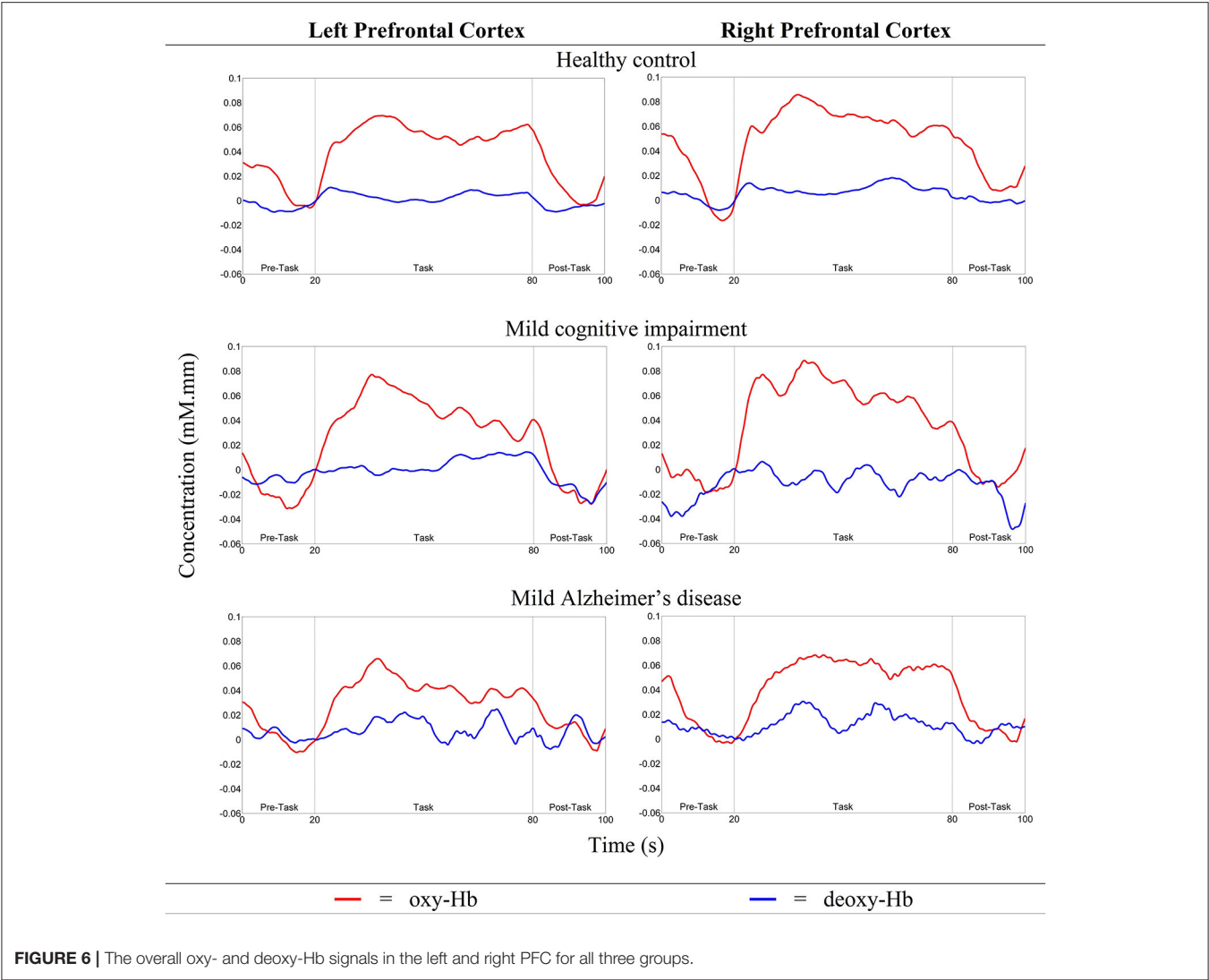


TABLE 2 | Pairwise correlation and simple linear regression results between fNIRS parameters, and MMSE score.

| | Left PFC | | Right PFC | |
|---|----------|----------|-----------|----------|
| | <i>m</i> | <i>R</i> | <i>m</i> | <i>R</i> |
| NUMBER OF ACTIVATED MEASUREMENT CHANNELS, | | | | |
| <i>N</i>_{left PFC/right PFC} vs. MMSE SCORE | | | | |
| HC | 0.1919 | 0.0540 | 0.3868 | 0.1393 |
| MCI | −0.1154 | −0.0634 | 0.1731 | 0.1032 |
| Mild AD | 0.2517 | 0.2914 | 0.0412 | 0.0429 |
| oxy-Hb CONCENTRATION CHANGE DURING THE ACTIVATION PERIOD, | | | | |
| Δoxy – <i>Hb</i>_{left PFC/right PFC} vs. MMSE SCORE | | | | |
| HC | 0.0010 | 0.0478 | 0.0035 | 0.1467 |
| MCI | 0.0152 | 0.5886 | 0.0086 | 0.4144 |
| Mild AD | −0.0006 | −0.0732 | −0.0019 | −0.2258 |
| TIME TAKEN TO ACHIEVE 25% OF PEAK ACTIVATION LEVEL, | | | | |
| <i>t</i>_{left PFC/right PFC} vs. MMSE Score | | | | |
| HC | −0.8478 | −0.3394 | −0.8567 | −0.3713 |
| MCI | −0.4644 | −0.3340 | 0.4153 | 0.3143 |
| Mild AD | 0.0961 | 0.0729 | −0.3606 | −0.2733 |
| SLOPE IN THE FIRST 5 S AFTER TASK ONSET, | | | | |
| <i>m</i>_{left PFC/right PFC} vs. MMSE SCORE | | | | |
| HC | 0.0008 | 0.2040 | 0.0020 | 0.2926 |
| MCI | 0.0007 | 0.2028 | −0.0001 | −0.0394 |
| Mild AD | 0.0000 | −0.0026 | 0.0004 | 0.2663 |

m, slope of the linear regression; *R*, correlation coefficient between the two parameters; HC, healthy control; MCI, mild cognitive impairment; mild AD, mild Alzheimer's disease.

findings (Johnson et al., 2000; Arai et al., 2006; Sperling, 2007; Driscoll et al., 2009; Woodard et al., 2009). As the dorsolateral PFC has been suggested to be associated with compensatory mechanisms (Erickson et al., 2007), the brain region-specific hyperactivation and hypoactivation observed in MCI and mild AD, respectively, might indicate the presence of neural compensation in the former, and the inability to compensate in the latter (Prvulovic et al., 2005; Clement and Belleville, 2012). The differences in the level of activation across groups might also explain the differences in performance in SVFT, particularly in the left hemisphere (Fallgatter et al., 1997; Arai et al., 2006; Herrmann et al., 2008). However, contrary to current opinion, in the present study activation in the right PFC was greater compared to the left PFC, and was consistent between all the groups. This might be for the following reasons: (a) the activation of right PFC in monitoring cognitive processes during SVFT might implicate a greater role compared to the left PFC; (b) the difference might be due to methodological differences, including the differences in recruitment strategies, brain region investigated and variations in research design and measurements used to describe the outcome. More specifically, the variations in research design refer to different assessment tools, inclusion and exclusion criteria (Fallgatter et al., 1997; Yeung et al., 2016), the numbers of groups (e.g., between normal aging and AD; Herrmann et al., 2008) or between normal aging and MCI (Yeung et al., 2016), and the focus on a brain region (e.g., frontal and parietal lobe). Similarly, differences in the outcomes measured

might exert a significant impact on the results, such as separate analyses of oxy-Hb and deoxy-Hb (Herrmann et al., 2008) and different baseline-corrected values (Fallgatter et al., 1997). The high variance observed here might be due to placement of the probes i.e., the position of the optodes relative to the skin (Strangman et al., 2003). Inter-subject anatomical variability, such as the thicknesses of the skull and the cerebrospinal fluid layers could have also caused the large variation in the fNIRS measurements (Okada and Delpy, 2003).

Moderate Positive Correlation

between Δ oxy – *Hb*_{left PFC} and MMSE Score in MCI

Finally, we propose an explanation for the moderately positive ($R = 0.5886$) linear relationship between the oxygenation level in the left PFC and the MMSE score, which was found only in MCI subjects, but not in HC and mild AD. This result reflects the fact that increasing activation actually contributed to the cognitive status in MCI, while both HC and mild AD did not show a similar trend. We suggest that in HC, the behavioral ceiling effect might have been achieved, as represented by the maximum score in MMSE (30) and this does not necessarily imply the presence of a parallel ceiling in brain activation (continuous increases in left PFC activation) (Hagenbeek et al., 2007). This might also indicate that the neural network is intact in HC. In addition, MMSE might be a relatively easy task for HC and the maximum score of 30 points may, therefore, not be a sensitive measure of cognitive status. MMSE is not designed to measure the cognitive ability of a healthy person. However, cognitive status in MCI is well below the ceiling. The pathogenesis of AD is characterized pathologically by brain accumulation of amyloid β -protein ($A\beta$) in the early stages (Jack et al., 2013) and $A\beta$ is thought to be the cause of neuronal dysfunction in AD (Palop and Mucke, 2010), which necessitates neural activity. Previous studies have reported that, relative to younger people or older people without $A\beta$, both cognitively normal older people with $A\beta$ deposition (Mormino et al., 2012) and MCI patients (Dickerson et al., 2004) exhibit higher neuronal activity during cognitive task performance. This phenomenon might be evidence of functional compensation keeping older people with $A\beta$ and MCI patients cognitively stable. In agreement with these results, the increase in oxygenation levels might represent an attempted compensatory response, and hence it is proportionate to the improvement in cognitive status in MCI. However, it is not possible to draw any conclusions with respect to the interpretation due to the small MCI sample size here. Meanwhile, for mild AD, no linear relationship was found between the oxygenation levels in the left PFC and the MMSE score. The degree of $A\beta$ deposition in the brain might reduce neural efficiency, which eventually causes progression to a more severe stage of AD (Landau et al., 2012). In such a situation, it is possible that patients eventually decline cognitively as any compensatory ability has been compromised (O'Brien et al., 2010). So, when patients progress into mild AD, their neural compensation ability might have been weakened or the neural networks might have been compromised to the point where higher oxygenation levels coupled with compensatory mechanisms are no longer enough to maintain cognitive status, unlike with MCI. Although

the explanations above might fit with the observations, it is important to note that various non-neural factors might confound such interpretations; disrupted neurovascular coupling that is associated with pathological conditions e.g., aging and disease (Buckner et al., 2000; D'Esposito et al., 2003) is one of the many factors. Other factors include alterations in perfusion and metabolism (El Fakhri et al., 2003), and vascular physiology (Mueggler et al., 2002). It has also been reported that employing different verbal memory strategies led to different patterns of cortical activation (Logan et al., 2002). A final factor that might have influenced PFC activation is the administration of medication e.g., cholinergic stimulation (Rombouts et al., 2002) and donepezil (McGeown et al., 2010).

Limitation

Subtypes of MCI need to be considered: a clinical presentation with memory impairment is characterized as amnesic MCI (aMCI), whereas the absence of memory impairment with the presence of impairment in one or more non-memory cognitive domains is characterized as non-amnesic MCI (naMCI). Furthermore, these subtypes can be further narrowed down into single and multi-domain impairments. It has been suggested that aMCI has a higher likelihood of progressing into AD, while naMCI is prone to developing into non-AD dementia (Petersen et al., 2009). Since the present study accessed a relatively small number of MCI patients, no attempt was made to exclude patients on the basis of other comorbidities. To substantiate the findings, research with a larger sample size might help ensure that participants with secondary comorbidities can be excluded. In addition, such a study could ensure that participants with different subtypes of MCI are assessed separately.

CONCLUSION

It was found that HC took a shorter time to achieve the targeted activation level in the left PFC compared to MCI and mild AD, while mild AD took a longer time than HC and MCI in the

right PFC. In addition, a steeper slope of activation was found in the right PFC of patients with MCI compared to HC and mild AD. The right PFC was particularly recruited in compensatory activity, which could be explained by the scaffolding theory of aging and cognition, and the inhibitory theory. Our results demonstrated, by using fNIRS, that compensation and neuroplasticity in the form of hyperactivation might be present in the PFC of MCI, but not in mild AD. Compensatory mechanisms might, therefore, have been compromised in mild AD. Time taken and the slope of activation were identified as key parameters of neuronal compensatory mechanisms, although the results presented here were not statistically significant after Bonferroni-Holm corrections. Future studies should look at these parameters individually. A moderately positive correlation between the oxygenation level in the left PFC and MMSE score was also found uniquely in MCI subjects. Longitudinal studies would be helpful in confirming whether task-elicited hyperactivation in MCI and hypoactivation in mild AD do indeed reflect the presence of compensatory mechanisms and the inability to compensate, respectively. If they do, future studies with a larger sample size could be directed toward investigating these fNIRS parameters as potential prognostic biomarkers of MCI and mild AD progression.

AUTHOR CONTRIBUTIONS

TT, EE, and SS designed the study. KY, WU, NN, PC, SC, and HY acquired the data. KY, WU, MK, and TT analyzed the data. KY, WU, and TT wrote the article, which all authors reviewed and approved for publication.

FUNDING

This work was partially supported by MARA [Grant no: MARA/UNI.1/33/06/02/15(4)] and the Ministry of Higher Education, Malaysia under Higher Institution Centre of Excellence (HiCOE) scheme to CISIR, UTP.

REFERENCES

- Alzheimer's Association (2016). 2016 Alzheimer's disease facts and figures. *Alzheimer's Dement.* 12, 459–509. doi: 10.1016/j.jalz.2016.03.001
- Arai, H., Takano, M., Miyakawa, K., Ota, T., Takahashi, T., Asaka, H., et al. (2006). A quantitative near-infrared spectroscopy study: a decrease in cerebral hemoglobin oxygenation in Alzheimer's disease and mild cognitive impairment. *Brain Cogn.* 61, 189–194. doi: 10.1016/j.bandc.2005.12.012
- Baazaoui, N., Flory, M., and Iqbal, K. (2017). Synaptic compensation as a probable cause of prolonged mild cognitive impairment in Alzheimer's disease: implications from a transgenic mouse model of the disease. *J. Alzheimers Dis.* 56, 1385–1401. doi: 10.3233/JAD-160845
- Banich, M. T. (1998). The missing link: the role of interhemispheric interaction in attentional processing. *Brain Cogn.* 36, 128–157. doi: 10.1006/brcg.1997.0950
- Barnes, J., Godbolt, A. K., Frost, C., Boyes, R. G., Jones, B. F., Schill, R. I., et al. (2007). Atrophy rates of the cingulate gyrus and hippocampus in AD and FTLD. *Neurobiol. Aging* 28, 20–28. doi: 10.1016/j.neurobiolaging.2005.11.012
- Buckner, R. L., Snyder, A. Z., Sanders, A. L., Raichle, M. E., and Morris, J. C. (2000). Functional brain imaging of young, nondemented, and demented older adults. *J. Cogn. Neurosci.* 12(Suppl. 2), 24–34. doi: 10.1162/089892900564046
- Cabeza, R. (2002). Hemispheric asymmetry reduction in older adults: the HAROLD model. *Psychol. Aging* 17, 85–100. doi: 10.1037/0882-7974.17.1.85
- Cabeza, R., Anderson, N. D., Locantore, J. K., and McIntosh, A. R. (2002). Aging gracefully: compensatory brain activity in high-performing older adults. *Neuroimage* 17, 1394–1402. doi: 10.1006/nimg.2002.1280
- Clement, F., and Belleville, S. (2012). Effect of disease severity on neural compensation of item and associative recognition in mild cognitive impairment. *J. Alzheimers Dis.* 29, 109–123. doi: 10.3233/JAD-2012-110426
- Clement, F., Gauthier, S., and Belleville, S. (2013). Executive functions in mild cognitive impairment: emergence and breakdown of neural plasticity. *Cortex* 49, 1268–1279. doi: 10.1016/j.cortex.2012.06.004
- Cotelli, M., Manenti, R., Cappa, S. F., Zanetti, O., and Miniussi, C. (2008). Transcranial magnetic stimulation improves naming in Alzheimer disease patients at different stages of cognitive decline. *Eur. J. Neurol.* 15, 1286–1292. doi: 10.1111/j.1468-1331.2008.02202.x
- Cox, S. R., Bastin, M. E., Ferguson, K. J., Allerhand, M., Royle, N. A., Maniega, S. M., et al. (2015). Compensation or inhibitory failure? Testing hypotheses of age-related right frontal lobe involvement in verbal memory ability using structural and diffusion MRI. *Cortex* 63, 4–15. doi: 10.1016/j.cortex.2014.08.001

- Cui, X. (2011). *Realtime Feedback from fNIRS Recording*. Available online at: <http://www.alivelearn.net/?p=1263>
- Cui, X., Bray, S., Bryant, D. M., Glover, G. H., and Reiss, A. L. (2011). A quantitative comparison of fNIRS and fMRI across multiple cognitive tasks. *Neuroimage* 54, 2808–2821. doi: 10.1016/j.neuroimage.2010.10.069
- Dannhauser, T. M., Shergill, S. S., Stevens, T., Lee, L., Seal, M., Walker, R. W., et al. (2008). An fMRI study of verbal episodic memory encoding in amnesic mild cognitive impairment. *Cortex* 44, 869–880. doi: 10.1016/j.cortex.2007.04.005
- de Chastelaine, M., Wang, T. H., Minton, B., Muftuler, L. T., and Rugg, M. D. (2011). The effects of age, memory performance, and callosal integrity on the neural correlates of successful associative encoding. *Cereb. Cortex* 21, 2166–2176. doi: 10.1093/cercor/bhq294
- D'Esposito, M., Deouell, L. Y., and Gazzaley, A. (2003). Alterations in the BOLD fMRI signal with ageing and disease: a challenge for neuroimaging. *Nat. Rev. Neurosci.* 4, 863–872. doi: 10.1038/nrn1246
- Dickerson, B. C., Salat, D. H., Bates, J. F., Atiya, M., Killiany, R. J., Greve, D. N., et al. (2004). Medial temporal lobe function and structure in mild cognitive impairment. *Ann. Neurol.* 56, 27–35. doi: 10.1002/ana.20163
- Driscoll, T., Kim, H.-T., Chae, B.-G., Kim, B.-J., Lee, Y.-W., Jokerst, N. M., et al. (2009). Memory metamaterials. *Science* 325, 1518–1521. doi: 10.1126/science.1176580
- Ehls, A.-C., Schneider, S., Dresler, T., and Fallgatter, A. J. (2014). Application of functional near-infrared spectroscopy in psychiatry. *Neuroimage* 85, 478–488. doi: 10.1016/j.neuroimage.2013.03.067
- El Fakhri, G., Kijewski, M. F., Johnson, K. A., Syrkina, G., Killiany, R. J., Becker, J. A., et al. (2003). MRI-guided SPECT perfusion measures and volumetric MRI in prodromal Alzheimer disease. *Arch. Neurol.* 60, 1066–1072. doi: 10.1001/archneur.60.8.1066
- Erickson, K. I., Colcombe, S. J., Wadhwa, R., Bherer, L., Peterson, M. S., Scalf, P. E., et al. (2007). Training-induced functional activation changes in dual-task processing: an fMRI study. *Cereb. Cortex* 17, 192–204. doi: 10.1093/cercor/bhj137
- Fallgatter, A., Roesler, M., Sitzmann, A., Heidrich, A., Mueller, T., and Strik, W. (1997). Loss of functional hemispheric asymmetry in Alzheimer's dementia assessed with near-infrared spectroscopy. *Cogn. Brain Res.* 6, 67–72. doi: 10.1016/S0926-6410(97)00016-5
- Ferrari, M., and Quaresima, V. (2012). A brief review on the history of human functional near-infrared spectroscopy (fNIRS) development and fields of application. *Neuroimage* 63, 921–935. doi: 10.1016/j.neuroimage.2012.03.049
- Folstein, M. F., Folstein, S. E., and McHugh, P. R. (1975). "Mini-mental state": a practical method for grading the cognitive state of patients for the clinician. *J. Psychiatr. Res.* 12, 189–198. doi: 10.1016/0022-3956(75)90026-6
- Gauthier, S., Reisberg, B., Zaudig, M., Petersen, R. C., Ritchie, K., Broich, K., et al. (2006). Mild cognitive impairment. *Lancet* 367, 1262–1270. doi: 10.1016/S0140-6736(06)68542-5
- Grossman, M., Smith, E. E., Koenig, P., Glosser, G., DeVita, C., Moore, P., et al. (2002). The neural basis for categorization in semantic memory. *Neuroimage* 17, 1549–1561. doi: 10.1006/nimg.2002.1273
- Gutierrez-Sigut, E., Payne, H., and MacSweeney, M. (2015). Investigating language lateralization during phonological and semantic fluency tasks using functional transcranial Doppler sonography. *Laterality* 20, 49–68. doi: 10.1080/1357650X.2014.914950
- Hagenbeek, R. E., Rombouts, S. A. R. B., Veltman, D. J., Van Strien, J. W., Witter, M. P., Scheltens, P., et al. (2007). Nonlinear changes in brain activity during continuous word repetition: an event-related multiparametric functional MRI imaging study. *Am. J. Neuroradiol.* 28, 1715–1721. doi: 10.3174/ajnr.A0632
- Hayama, H. R., and Rugg, M. D. (2009). Right dorsolateral prefrontal cortex is engaged during post-retrieval processing of both episodic and semantic information. *Neuropsychologia* 47, 2409–2416. doi: 10.1016/j.neuropsychologia.2009.04.010
- Heinzel, S., Hauesinger, F. B., Hahn, T., Ehls, A.-C., Plichta, M. M., and Fallgatter, A. J. (2013). Variability of (functional) hemodynamics as measured with simultaneous fNIRS and fMRI during intertemporal choice. *Neuroimage* 71, 125–134. doi: 10.1016/j.neuroimage.2012.12.074
- Herrmann, M. J., Langer, J. B., Jacob, C., Ehls, A.-C., and Fallgatter, A. J. (2008). Reduced prefrontal oxygenation in Alzheimer disease during verbal fluency tasks. *Am. J. Geriatr. Psychiatry* 16, 125–135. doi: 10.1097/JGP.0b013e3180cc1fbc
- Ishii-Takahashi, A., Takizawa, R., Nishimura, Y., Kawakubo, Y., Kuwabara, H., Matsubayashi, J., et al. (2014). Prefrontal activation during inhibitory control measured by near-infrared spectroscopy for differentiating between autism spectrum disorders and attention deficit hyperactivity disorder in adults. *Neuroimage Clin.* 4, 53–63. doi: 10.1016/j.nicl.2013.10.002
- Jack, C. R. Jr., Knopman, D. S., Jagust, W. J., Petersen, R. C., Weiner, M. W., Aisen, P. S., et al. (2013). Tracking pathophysiological processes in Alzheimer's disease: an updated hypothetical model of dynamic biomarkers. *Lancet Neurol.* 12, 207–216. doi: 10.1016/S1474-4422(12)70291-0
- Joanette, Y., and Goulet, P. (1986). Criterion-specific reduction of verbal fluency in right brain-damaged right-handers. *Neuropsychologia* 24, 875–879. doi: 10.1016/0028-3932(86)90087-4
- Johnson, S. C., Saykin, A. J., Baxter, L. C., Flashman, L. A., Santulli, R. B., McAllister, T. W., et al. (2000). The relationship between fMRI activation and cerebral atrophy: comparison of normal aging and Alzheimer disease. *Neuroimage* 11, 179–187. doi: 10.1006/nimg.1999.0530
- Johnson, S. C., Schmitz, T. W., Moritz, C. H., Meyerand, M. E., Rowley, H. A., Alexander, A. L., et al. (2006). Activation of brain regions vulnerable to Alzheimer's disease: the effect of mild cognitive impairment. *Neurobiol. Aging* 27, 1604–1612. doi: 10.1016/j.neurobiolaging.2005.09.017
- Karantzoulis, S., and Galvin, J. E. (2014). Distinguishing Alzheimer's disease from other major forms of dementia. *Expert Rev. Neurother.* 11, 1579–1591. doi: 10.1586/ern.11.155
- Klem, G. H., Luders, H. O., Jasper, H. H., and Elger, C. (1999). The ten-twenty electrode system of the International Federation. The international federation of clinical neurophysiology. *Electroencephalogr. Clin. Neurophysiol. Suppl.* 52, 3–6.
- Lambon Ralph, M. A., Graham, K. S., Ellis, A. W., and Hodges, J. R. (1998). Naming in semantic dementia—what matters? *Neuropsychologia* 36, 775–784. doi: 10.1016/S0028-3932(97)00169-3
- Landau, S. M., Mintun, M. A., Joshi, A. D., Koeppe, R. A., Petersen, R. C., Aisen, P. S., et al. (2012). Amyloid deposition, hypometabolism, and longitudinal cognitive decline. *Ann. Neurol.* 72, 578–586. doi: 10.1002/ana.23650
- Li, T., Gong, H., and Luo, Q. (2011). Visualization of light propagation in visible Chinese human head for functional near-infrared spectroscopy. *J. Biomed. Opt.* 16, 045001. doi: 10.1117/1.3567085
- Logan, J. M., Sanders, A. L., Snyder, A. Z., Morris, J. C., and Buckner, R. L. (2002). Under-recruitment and nonselective recruitment: dissociable neural mechanisms associated with aging. *Neuron* 33, 827–840. doi: 10.1016/S0896-6273(02)00612-8
- Luu, S., and Chau, T. (2009). Decoding subjective preference from single-trial near-infrared spectroscopy signals. *J. Neural Eng.* 6:016003. doi: 10.1088/1741-2560/6/1/016003
- Magni, E., Binetti, G., Bianchetti, A., Rozzini, R., and Trabucchi, M. (1996). Mini-Mental State Examination: a normative study in Italian elderly population. *Eur. J. Neurol.* 3, 198–202. doi: 10.1111/j.1468-1331.1996.tb00423.x
- Maki, A., Yamashita, Y., Ito, Y., Watanabe, E., Mayanagi, Y., and Koizumi, H. (1995). Spatial and temporal analysis of human motor activity using noninvasive NIR topography. *Med. Phys.* 22, 1997–2005. doi: 10.1118/1.597496
- Mandzia, J. L., McAndrews, M. P., Grady, C. L., Graham, S. J., and Black, S. E. (2009). Neural correlates of incidental memory in mild cognitive impairment: an fMRI study. *Neurobiol. Aging* 30, 717–730. doi: 10.1016/j.neurobiolaging.2007.08.024
- Marner, L., Nyengaard, J. R., Tang, Y., and Pakkenberg, B. (2003). Marked loss of myelinated nerve fibers in the human brain with age. *J. Comp. Neurol.* 462, 144–152. doi: 10.1002/cne.10714
- McGeown, W. J., Shanks, M. F., Forbes-McKay, K. E., Waiter, G. D., Elrick, I., Venneri, M. G., et al. (2010). Established donepezil treatment modulates task relevant regional brain activation in early Alzheimer's disease. *Curr. Alzheimer Res.* 7, 415–427. doi: 10.2174/156720510791383877
- McNab, F., and Klingberg, T. (2008). Prefrontal cortex and basal ganglia control access to working memory. *Nat. Neurosci.* 11, 103–107. doi: 10.1038/nn2024
- Mormino, E. C., Brandel, M. G., Madison, C. M., Marks, S., Baker, S. L., and Jagust, W. J. (2012). Aβ deposition in aging is associated with increases in brain activation during successful memory encoding. *Cereb. Cortex* 22, 1813–1823. doi: 10.1093/cercor/bhr255
- Morris, J. C. (1993). The Clinical Dementia Rating (CDR): current version and scoring rules. *Neurology* 43, 2412–2414. doi: 10.1212/WNL.43.11.2412-a

- Moss, H. E., Tyler, L. K., Devlin, J. T., Forde, E., and Humphreys, G. (2002). "The emergence of category-specific deficits in a distributed semantic system," *Category-Specificity in Brain and Mind*, eds E. Forde and G. Humphreys (Hove: Psychology Press), 115–148.
- Mueggler, T., Sturchler-Pierrat, C., Baumann, D., Rausch, M., Staufenbiel, M., and Rudin, M. (2002). Compromised hemodynamic response in amyloid precursor protein transgenic mice. *J. Neurosci.* 22, 7218–7224.
- Nho, K., Risacher, S. L., Crane, P. K., DeCarli, C., Glymour, M. M., Habeck, C., et al. (2012). Voxel and surface-based topography of memory and executive deficits in mild cognitive impairment and Alzheimer's disease. *Brain Imaging Behav.* 6, 551–567. doi: 10.1007/s11682-012-9203-2
- O'Brien, J. L., O'Keefe, K. M., LaViolette, P. S., DeLuca, A. N., Blacker, D., Dickerson, B. C., et al. (2010). Longitudinal fMRI in elderly reveals loss of hippocampal activation with clinical decline. *Neurology* 74, 1969–1976. doi: 10.1212/WNL.0b013e3181e3966e
- Okada, E., and Delpy, D. T. (2003). Near-infrared light propagation in an adult head model. II. Effect of superficial tissue thickness on the sensitivity of the near-infrared spectroscopy signal. *Appl. Opt.* 42, 2915–2922. doi: 10.1364/AO.42.002915
- Palop, J. J., and Mucke, L. (2010). Amyloid- β induced neuronal dysfunction in Alzheimer's disease: from synapses toward neural networks. *Nat. Neurosci.* 13, 812–818. doi: 10.1038/nn.2583
- Park, D. C., and Bischof, G. N. (2013). The aging mind: neuroplasticity in response to cognitive training. *Dialogues Clin. Neurosci.* 15, 109–119.
- Patterson, K., and Hodges, J. R. (1992). Deterioration of word meaning: implications for reading. *Neuropsychologia* 30, 1025–1040. doi: 10.1016/0028-3932(92)90096-5
- Perry, R. J., Watson, P., and Hodges, J. R. (2000). The nature and staging of attention dysfunction in early (minimal and mild) Alzheimer's disease: relationship to episodic and semantic memory impairment. *Neuropsychologia* 38, 252–271. doi: 10.1016/S0028-3932(99)00079-2
- Petersen, R. C. (2009). Early diagnosis of Alzheimer's disease: is MCI too late? *Curr. Alzheimer Res.* 6, 324–330. doi: 10.2174/156720509788929237
- Petersen, R. C., Roberts, R. O., Knopman, D. S., Boeve, B. F., Geda, Y. E., Ivnik, R. J., et al. (2009). Mild cognitive impairment: ten years later. *Arch. Neurol.* 66, 1447–1455. doi: 10.1001/archneur.2009.266
- Price, C. J., and Friston, K. J. (2002). Degeneracy and cognitive anatomy. *Trends Cogn. Sci.* 6, 416–421. doi: 10.1016/S1364-6613(02)01976-9
- Prvulovic, D., Van de Ven, V., Sack, A. T., Maurer, K., and Linden, D. E. (2005). Functional activation imaging in aging and dementia. *Psychiatry Res.* 140, 97–113. doi: 10.1016/j.psychres.2005.06.006
- Ranger, M., Johnston, C. C., Limperopoulos, C., Rennick, J. E., and du Plessis, A. J. (2011). Cerebral near-infrared spectroscopy as a measure of nociceptive evoked activity in critically ill infants. *Pain Res. Manag.* 16, 331–336. doi: 10.1155/2011/891548
- Reuter-Lorenz, P. A., Jonides, J., Smith, E. E., Hartley, A., Miller, A., Marshuetz, C., et al. (2000). Age differences in the frontal lateralization of verbal and spatial working memory revealed by PET. *J. Cogn. Neurosci.* 12, 174–187. doi: 10.1162/089892900561814
- Rockwood, K., Strang, D., MacKnight, C., Downer, R., and Morris, J. C. (2000). Interrater reliability of the Clinical Dementia Rating in a multicenter trial. *J. Am. Geriatr. Soc.* 48, 558–559. doi: 10.1111/j.1532-5415.2000.tb05004.x
- Rombouts, S., Barkhof, F., van Meel, C. S., and Scheltens, P. (2002). Alterations in brain activation during cholinergic enhancement with rivastigmine in Alzheimer's disease. *J. Neurol. Neurosurg. Psychiatry* 73, 665–671. doi: 10.1136/jnnp.73.6.665
- Salat, D. H., Kaye, J. A., and Janowsky, J. S. (2001). Selective preservation and degeneration within the prefrontal cortex in aging and Alzheimer disease. *Arch. Neurol.* 58, 1403–1408. doi: 10.1001/archneur.58.9.1403
- Sato, H., Kiguchi, M., Maki, A., Fuchino, Y., Obata, A., Yoro, T., et al. (2006). Within-subject reproducibility of near-infrared spectroscopy signals in sensorimotor activation after 6 months. *J. Biomed. Opt.* 11, 014021. doi: 10.1117/1.2166632
- Scarmeas, N., Zarahn, E., Anderson, K. E., Hilton, J., Flynn, J., Van Heertum, R. L., et al. (2003). Cognitive reserve modulates functional brain responses during memory tasks: a PET study in healthy young and elderly subjects. *Neuroimage* 19, 1215–1227. doi: 10.1016/S1053-8119(03)00074-0
- Shi, F., Liu, B., Zhou, Y., Yu, C., and Jiang, T. (2009). Hippocampal volume and asymmetry in mild cognitive impairment and Alzheimer's disease: meta-analyses of MRI studies. *Hippocampus* 19, 1055–1064. doi: 10.1002/hipo.20573
- Sperling, R. (2007). Functional MRI studies of associative encoding in normal aging, mild cognitive impairment, and Alzheimer's disease. *Ann. N. Y. Acad. Sci.* 1097, 146–155. doi: 10.1196/annals.1379.009
- Strangman, G. E., Li, Z., and Zhang, Q. (2013). Depth sensitivity and source-detector separations for near infrared spectroscopy based on the Colin27 brain template. *PLoS ONE* 8:e66319. doi: 10.1371/journal.pone.0066319
- Strangman, G., Franceschini, M. A., and Boas, D. A. (2003). Factors affecting the accuracy of near-infrared spectroscopy concentration calculations for focal changes in oxygenation parameters. *Neuroimage* 18, 865–879. doi: 10.1016/S1053-8119(03)00021-1
- Sutoko, S., Sato, H., Maki, A., Kiguchi, M., Hirabayashi, Y., Atsumori, H., et al. (2016). Tutorial on platform for optical topography analysis tools. *Neurophotonics* 3:010801. doi: 10.1117/1.NPh.3.1.010801
- Takizawa, R., Fukuda, M., Kawasaki, S., Kasai, K., Mimura, M., Pu, S., et al. (2014). Neuroimaging-aided differential diagnosis of the depressive state. *Neuroimage* 85, 498–507. doi: 10.1016/j.neuroimage.2013.05.126
- Thompson-Schill, S. L., D'Esposito, M., Aguirre, G. K., and Farah, M. J. (1997). Role of left inferior prefrontal cortex in retrieval of semantic knowledge: a reevaluation. *Proc. Natl. Acad. Sci. U.S.A.* 94, 14792–14797. doi: 10.1073/pnas.94.26.14792
- Wagner, S., Sebastian, A., Lieb, K., Tüscher, O., and Tadić, A. (2014). A coordinate-based ALE functional MRI meta-analysis of brain activation during verbal fluency tasks in healthy control subjects. *BMC Neurosci.* 15:19. doi: 10.1186/1471-2202-15-19
- Woodard, J. L., Seidenberg, M., Nielson, K. A., Antuono, P., Guidotti, L., Durgerian, S., et al. (2009). Semantic memory activation in amnesic mild cognitive impairment. *Brain* 132, 2068–2078. doi: 10.1093/brain/awp157
- Yankner, B. A., Lu, T., and Loerch, P. (2008). The aging brain. *Annu. Rev. Pathol.* 3, 41–66. doi: 10.1146/annurev.pathmechdis.2.010506.092044
- Yeung, M. K., Sze, S. L., Woo, J., Kwok, T., Shum, D. H., Yu, R., et al. (2016). Altered frontal lateralization underlies the category fluency deficits in older adults with mild cognitive impairment: a near-infrared spectroscopy study. *Front. Aging Neurosci.* 8:59. doi: 10.3389/fnagi.2016.00059
- Zanto, T. P., Rubens, M. T., Thangavel, A., and Gazzaley, A. (2011). Causal role of the prefrontal cortex in top-down modulation of visual processing and working memory. *Nat. Neurosci.* 14, 656–661. doi: 10.1038/nn.2773

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.

Copyright © 2017 Yap, Ung, Ebenezer, Nordin, Chin, Sugathan, Chan, Yip, Kiguchi and Tang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



White Matter Tract Integrity in Alzheimer's Disease vs. Late Onset Bipolar Disorder and Its Correlation with Systemic Inflammation and Oxidative Stress Biomarkers

Ariadna Besga^{1,2*}, Darya Chyzhyk^{3,4}, Itxaso Gonzalez-Ortega^{5,6}, Jon Echeveste⁷, Marina Graña-Lecuona⁴, Manuel Graña^{3,4} and Ana Gonzalez-Pinto^{1,5,8}

¹ Centre for Biomedical Research Network on Mental Health, Spain, ² Department of Internal Medicine of Hospital Universitario de Alava, Vitoria, Spain, ³ Computational Intelligence Group, University of the Basque Country (UPV/EHU), San Sebastian, Spain, ⁴ ACPySS, San Sebastian, Spain, ⁵ Department of Psychiatry, University Hospital of Alava-Santiago, Vitoria, Spain, ⁶ School of Psychology, University of the Basque Country (UPV/EHU), San Sebastian, Spain, ⁷ Magnetic Resonance Imaging Department, Osatek, Vitoria, Spain, ⁸ School of Medicine, University of the Basque Country (UPV/EHU), Vitoria, Spain

OPEN ACCESS

Edited by:

Juan Manuel Gorriz,
University of Granada, Spain

Reviewed by:

Abdelbasset Brahim,
University of Orléans, France
Dario Maravall,
Universidad Politécnica de Madrid
(UPM), Spain
Lingzhong Fan,
Institute of Automation (CAS), China

*Correspondence:

Ariadna Besga
ariadna.besgabasterra@osakidetza.net

Received: 12 February 2017

Accepted: 23 May 2017

Published: 16 June 2017

Citation:

Besga A, Chyzhyk D, Gonzalez-Ortega I, Echeveste J, Graña-Lecuona M, Graña M and Gonzalez-Pinto A (2017) White Matter Tract Integrity in Alzheimer's Disease vs. Late Onset Bipolar Disorder and Its Correlation with Systemic Inflammation and Oxidative Stress Biomarkers. *Front. Aging Neurosci.* 9:179. doi: 10.3389/fnagi.2017.00179

Background: Late Onset Bipolar Disorder (LOBD) is the development of Bipolar Disorder (BD) at an age above 50 years old. It is often difficult to differentiate from other aging dementias, such as Alzheimer's Disease (AD), because they share cognitive and behavioral impairment symptoms.

Objectives: We look for WM tract voxel clusters showing significant differences when comparing of AD vs. LOBD, and its correlations with systemic blood plasma biomarkers (inflammatory, neurotrophic factors, and oxidative stress).

Materials: A sample of healthy controls (HC) ($n = 19$), AD patients ($n = 35$), and LOBD patients ($n = 24$) was recruited at the Alava University Hospital. Blood plasma samples were obtained at recruitment time and analyzed to extract the inflammatory, oxidative stress, and neurotrophic factors. Several modalities of MRI were acquired for each subject,

Methods: Fractional anisotropy (FA) coefficients are obtained from diffusion weighted imaging (DWI). Tract based spatial statistics (TBSS) finds FA skeleton clusters of WM tract voxels showing significant differences for all possible contrasts between HC, AD, and LOBD. An ANOVA F -test over all contrasts is carried out. Results of F -test are used to mask TBSS detected clusters for the AD > LOBD and LOBD > AD contrast to select the image clusters used for correlation analysis. Finally, Pearson's correlation coefficients between FA values at cluster sites and systemic blood plasma biomarker values are computed.

Results: The TBSS contrasts with by ANOVA F -test has identified strongly significant clusters in the forceps minor, inferior longitudinal fasciculus, inferior fronto-occipital fasciculus, and cingulum gyrus. The correlation analysis of these tract clusters found strong negative correlation of AD with the nerve growth factor (NGF) and brain derived

neurotrophic factor (BDNF) blood biomarkers. Negative correlation of AD and positive correlation of LOBD with inflammation biomarker IL6 was also found.

Conclusion: TBSS voxel clusters tract atlas localizations are consistent with greater behavioral impairment and mood disorders in LOBD than in AD. Correlation analysis confirms that neurotrophic factors (i.e., NGF, BDNF) play a great role in AD while are absent in LOBD pathophysiology. Also, correlation results of IL1 and IL6 suggest stronger inflammatory effects in LOBD than in AD.

Keywords: late onset bipolar disorder, tract based spatial statistics, Alzheimer disease, inflammatory biomarkers, multimodal brain data analysis, nerve growth factors

INTRODUCTION

Bipolar disorder (BD) is a chronic mood disorder characterized by manic and depressive alternating episodes, interspersed by euthymic periods. Age of onset may be determined by environmental and genetic conditions (Bauer M. et al., 2014; Martinez-Cengotitabengoa et al., 2014; Bauer et al., 2015a,b). Commonly, BD onset happens during youth years, leading to cognitive, affective, and functional impairment (Forcada et al., 2015). When the onset age is above 50 years, it is considered a late onset BD (LOBD) (Depp and Jeste, 2004; Zanetti et al., 2007; Prabhakar and Balon, 2010; Besga et al., 2011; Carlino et al., 2013; Chou et al., 2015), which may be difficult to differentiate from Alzheimer's disease (AD), because of overlapping symptoms (Zahodne et al., 2015). Another example of the fuzzy boundaries between brain pathologies is the discovery of an AD biomarker signature that also identifies Parkinson's Disease patients with dementia (PDD) (Berlyand et al., 2016) opening the door for crossover treatment of PDD with AD therapies. This trend is appreciated in recent studies comparing BD and AD patients (Berridge, 2013). Specifically, inflammation and oxidative stress biomarkers have been identified for AD (Akiyama et al., 2000; Kamer et al., 2008; Sardi et al., 2011), LOBD (Goldstein et al., 2009; Konradi et al., 2012; Leboyer et al., 2012; Lee et al., 2013; Bauer I. E. et al., 2014; Hope et al., 2015), depression, and mania (Brydon et al., 2009; Dickerson et al., 2013; Castanon et al., 2014; Singhal et al., 2014). Common traits between LOBD and AD are described in Besga et al. (2015). Common psychiatric symptoms in AD which are shared with the profile observed in LOBD patients are: agitation, euphoria, disinhibition, over-activity without agitation, aggression, affective lability, dysphoria, apathy, impaired self-regulation, and psychosis (Albert and Blacker, 2006; Zahodne et al., 2015).

This paper contains a new contribution to a comparative study of AD vs. LOBD patients that has been carried out for some time. In this study, demographic and other data gathered from the patients at recruitment, such as psychological tests and MRI data, has been described in Besga et al. (2012), Graña et al. (2011), and Besga et al. (2015, 2016), therefore description of materials can not be duplicated here without breaking imposed journal self-plagiarism rules. Consequently we refer the reader to these publications, while here we provide a summary account of the study and results achieved and reported in previous publications. Over one hundred subjects older than 64 years were

recruited, including healthy controls (HC), and AD and LOBD patients. These subjects were treated to neuropsychological tests, blood extraction for plasma biomarkers measurement, and the acquisition of several modalities of magnetic resonance imaging (MRI). Specifically, Diffusion-Weighted Imaging (DWI) was acquired in order to study significant differences in the white matter (WM) structure. Reasons for eligibility and discarding of patients and full account of the materials are given in Besga et al. (2012), Graña et al. (2011), and Besga et al. (2015, 2016), and we dare not reproduced them here. Previously reported results of this study have been the following ones:

1. We demonstrate good discrimination between AD and LOBD populations using whole brain fractional anisotropy (FA) coefficients extracted from the DWI data (Graña et al., 2011; Besga et al., 2012), using multivariate machine learning for computer aided diagnosis (CAD) system design (Sigut et al., 2007; Salas-Gonzalez et al., 2009; Ramirez et al., 2010; Savio et al., 2011; Westman et al., 2011; Termenon et al., 2013). Though the classification performance results were good, the localization of effects in brain regions was not as satisfactory due to the feature extraction process.
2. We achieved encouraging classification results based on the clinical, neuropsychological test, and blood plasma biomarkers (a subset of the ones used in the this paper Besga et al., 2015) using machine learning techniques. We found that clinical variables have the greatest discriminant power. Blood plasma biomarkers alone have little discriminant power but help improve the clinical variables. Besides, we had no anatomical correspondences of the findings because no imaging data was involved.
3. Looking for inferences about the anatomical correlations of blood biomarkers we applied canonical correlation between them and whole brain FA data (Besga et al., 2016) using eigenanatomy tools that decompose the FA volume into eigenvolumes maximally correlated with plasma biomarkers (Avants et al., 2012, 2014) for feature extraction and ensuing AD vs. LOBD classification by machine learning. We found positive correlations of the oxido-nitrosative stress biomarker malondialdehyde (MDA) (Besga et al., 2016) with voxel clusters in the superior corona radiata, internal capsule, and superior longitudinal fasciculus. However, confirmation by classification performance was moderate, accuracy was below 80%.

Besides the commented limitations of previous works motivating further research, measurements of two new inflammation biomarkers were made available for new computational explorations.

Contributions of This Paper

The process carried out in this paper has two phases: first, cluster selection in FA volumes, and, second, a correlation analysis between image clusters and blood biomarkers. For cluster selection, we apply Tract-Based Spatial Statistics (TBSS) (Smith et al., 2006; Bach et al., 2014) to find significant WM tract differences between AD and LOBD patients. We expect TBSS to provide tract specific effects, improving the anatomical finding reported by our previous approaches (Graña et al., 2011; Besga et al., 2012, 2016). TBSS identifies WM tract voxel clusters with significant difference in FA between AD and LOBD on the mean FA skeleton. To enhance localization power we carry out an ANOVA *F*-test. The correlation analysis is carried out by computing Pearson's correlation coefficients of the FA values at voxel sites and blood biomarker values across all subjects. Machine learning is not used because of the small sample size that makes cross-validation results very unstable and not significant.

METHODS

Ethics Statement

The ethics committee of the Alava University Hospital, Spain, approved this study. All patients gave their written consent to participate in the study, which was conducted according to the provisions of the Helsinki declaration. After written informed consent was obtained, venous blood samples (10 mL) were collected from the volunteers, after which all the mood scales and cognitive tests were performed. The study has been registered as an observation trial¹ in the ISRCTN registry.

Blood Plasma Biomarkers

The blood plasma biomarkers selected for analysis include:

- Neurotrophins: nerve growth factor (NGF) and brain-derived neurotrophic factor (BDNF).
- Inflammation biomarkers (Akiyama et al., 2000; Kamer et al., 2008; Goldstein et al., 2009; Sardi et al., 2011; Lee et al., 2013; Garcia-Bueno and et al., 2014): Cytokines Interleukins 1 and 6 (IL-1 β , IL-6) and Tumor Necrosis Factor (TNF α), and the Cyclooxygenases (COX-1 and COX-2) by-products Prostaglandin E₂ (PGE2) and 15d-Prostaglandin J₂ (PGJ2).
- Oxidative stress biomarkers: nitrites (NO₂) and malondialdehyde (MDA).

Their measurements are described in Besga et al. (2015), except for PGE2 and PGJ2 which were measured by enzyme immunoassay (EIA) using reagents in kit form (Prostaglandin E2 EIA Kit-Monoclonal; Cayman Chemical Europe, Tallinn, Estonia and 15-deoxy- Δ 12,14-Prostaglandin J2 ELISA Kit DRG Diagnostics, Marburg, Germany). Samples were measured following manufacturer's instructions.

Diffusion Weighted Imaging

Diffusion-Weighted Imaging (DWI) uses MRI acquisition sequences computing signal differences along several gradient directions in order to obtain a signal that measures water diffusion. Diffusion Tensor Imaging (DTI) is a compact representation by means of 3×3 matrix of water diffusion in each spatial direction at each voxel (Basser et al., 1994; Pierpaoli et al., 1996). Specifically, in this paper we will work on FA values, which are computed as follows:

$$FA(j) = \sqrt{\frac{3 \sum_{i=1}^3 (\lambda_i - \bar{\lambda})^2}{2 \sum_{i=1}^3 \lambda_i^2}}, \quad (1)$$

where $\{\lambda_1, \lambda_2, \lambda_3\}$ are the diffusion tensor eigenvalues at the voxel site j . The specific parameters of the data capture on a 1.5 Tesla scanner (Magnetom Avanto, Siemens), data preprocessing, computing FA, and registration have been given in Besga et al. (2016). We use the FSL software suite (<http://www.fmrib.ox.ac.uk/fsl/>) to carry out DWI preprocessing, DTI estimation (Behrens et al., 2003), image registration (Andersson et al., 2007a,b), and TBSS described below. We did not perform spatial smoothing. The pre-processing consists in the removal of non-brain voxels using the brain extraction tool (BET) from FSL, the correction of eddy currents artifacts, and the rigid registration of the gradient images to cope with motion of the subject. On the spatially aligned DWI we estimate the diffusion tensor at each voxel, and the FA values. The FA volumes are then spatially normalized by non-linear registration to the FMRIB58_FA template provided with FSL standards. We have not computed a template from the actual FA dataset because the population is small and very heterogenous so the resulting mean FA template is quite noisy and blurry. Besides we need to register the data to the MNI152 space in order to report atlas based localizations, which is already done in the template provided by FSL. We do not carry out any intensity normalization on the FA images.

TBSS

We apply TBSS (Smith et al., 2006; Bach et al., 2014), a module of FSL (Smith et al., 2004)² to detect differences in white matter tracts between HC, AD, and LOBD subjects. The specific TBSS procedure applied is as follows: (1) We warp the FA volumes according to the registration carried out before, so we have all aligned in the common space. (2) We compute the mean FA image and extract the common skeleton from it by morphological image processing. This skeleton is assumed to represent the centerline of the WM tracts in all the FA volumes. Each subject's aligned FA data is then projected onto this skeleton. This projection is achieved by assigning the closest local maximum of FA value in the orthogonal direction to the skeleton. (3) For each possible contrast (i.e., HC > AD, AD > LOBD, HC > LOBD, AD > HC, LOBD > AD, LOBD > HC, and ANOVA *F*-test over all pairwise contrasts) we compute a permutation test applying the randomize tool of FSL with 50,000 permutations and threshold free cluster enhancement (TFCE)(Smith and Nichols,

¹<http://www.controlled-trials.com/search?q=HS%2FPI2010001>

²<http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/TBSS>

2009) skeleton cluster selection over the contrast statistics. The advantage of TFCE over other methods is that it combines the spatial and statistics value optimally achieving high significance, and its parameters are already optimized by an automated procedure.

Biomarker Correlation Analysis

The correlation analysis is applied to the clusters selected by each binary contrast masked by the *F*-test cluster detection. The resulting clusters are much smaller than the binary contrast detection but more specific. Considering independently each voxel site *j* of the selected clusters, we build a vector v_j composed of the FA intensities at the *j*-th voxel site across all the subjects. We compute Pearson's correlation coefficient between this vector and the value of the blood biomarker for this subject, denoted y_i in the following, obtaining the correlation values at each voxel site. Pearson's correlation (Pearson, 1895; Kendall and Stuart, 1973) at the *j*-th voxel site is computed as follows:

$$r_{v_j, y} = \frac{n \sum_i v_{ij} y_i - \sum_i v_{ij} \sum_i y_i}{\sqrt{n \sum_i v_{ij}^2 - (\sum_i v_{ij})^2} \sqrt{n \sum_i y_i^2 - (\sum_i y_i)^2}}, \quad (2)$$

where v_{ij} is the value of FA at the *j*-th voxel site in the *i*-th subject, and y_i is the blood plasma biomarker value of that *i*-th subject. We compute correlation values for each subpopulation (i.e., AD and LOBD) independently (following a reviewer recommendation), retaining voxels significantly correlated ($p < 0.01$) for examination.

Atlas Based Effect Localization

Locations reported by atlasquery tool from FSL using the JHU White-Matter Tractography Atlas are collected for each contrast after the permutation test, and each correlation analysis between detected FA skeleton clusters and blood biomarkers. The report produced by the atlasquery tool are the probability of a cluster voxel to belong to a tract, therefore the size of the cluster falling in a tract is computed as product of atlasquery probabilities and the total detection volume.

RESULTS

Figure 1 plots the size (logarithmic scale) of the skeleton clusters found by the permutation test for each of the contrasts including the *F*-test for each tract that can be identified in the JHU White-Matter Tractography Atlas, i.e., Right (R) and left (L) hemispheres of anterior thalamic radiation (ATR), corticospinal tract (CT), cingulum (cingulate gyrus) (C_CG), Cingulum (hippocampus) (CH), forceps minor (Fmi), forceps major (Fma), inferior fronto-occipital fasciculus (IFOF), inferior longitudinal fasciculus (ILF), superior longitudinal fasciculus (SFL), uncinate fasciculus (UF), temporal part of SLF (SLFT). Before masking with *F*-test results, pairwise contrasts greatest effects are located in the CT, ATR, Fma, and Fmi. However big, CT clusters are of similar size for all contrast, and disappear after *F*-test masking. Similarly, ATR clusters are reduced while ILF clusters relative importance increase after *F*-test masking. Clusters where HC or

LOBD have greater FA signal than AD (HC > AD, LOBD > AD) are much bigger (note log scale in the plot) across all the tracts than the converse (AD > HC, AD > LOBD). Also, clusters of HC > LOBD are much bigger than LOBD > HC. These differences in effect size are strongly significant ($p < 0.00001$, pairwise *t*-tests). The *F*-test selection involves mostly the C_CG, Fmi, IFOF, and ILF tracts. *F*-test detections in other tract is marginal, though we have included them in the correlation analysis.

Figure 2 illustrates the clusters detected by the AD > LOBD and LOBD > AD contrasts after masking with the *F*-test clusters. **Figure 2A** presents the mean FA volume and its skeleton (green). **Figure 2B** presents the *F*-test statistics (red). **Figure 2C** presents the significant clusters of the AD > LOBD contrast (blue). **Figure 2D** presents significant clusters of the LOBD > AD contrast.

Figure 3 shows in graphical form the correlation analysis results for each contrast of interest (AD > LOBD, LOBD > AD) and population (AD, LOBD). Nodes in the graph are either blood biomarkers (blue ellipsoids) or white matter tracts (rectangles), red arrows denote negative correlation, green arrows denote positive correlation. **Table 1** gives the sizes of the clusters of positive and negative correlations. Notice that the AD > LOBD contrast effects are very small according to **Figure 1**. The strongest effects correspond to neurotrophic factors BDNF, and NGF, and inflammation marker IL6. The NGF accounts for 70% of all the correlation effects. BDNF appears positively correlated to LOBD and negatively correlated to AD.

DISCUSSION

We study differential effects in the WM tracts and their correlation to plasma biomarkers looking for new insights into the pathophysiological processes underlying AD and LOBD (Lebert et al., 2008; Carlino et al., 2013; Grande et al., 2014). We know that cognitive degradation in LOBD is a key factor in differential diagnosis between LOBD and AD. Besides cognitive performance, behavioral disorders are also closely related to the overall functionality of the patients. Agitation, euphoria and disinhibition are the non-cognitive neuropsychological variables having the greatest discrimination power in the classification of patients into AD or LOBD (Besga et al., 2015), while memory cognitive domain performance is essential in clinical practice for the detection and diagnosis of AD (Weintraub et al., 2012). Besides, recent studies have revealed that significant cognitive impairment in BD compared to controls may allow to discriminate type I and II BD patients (Arahamian et al., 2014; Sparding et al., 2015), and may affect its prognosis, as it happens in patients with dementia (Kawas et al., 2003). Previously, some authors suggested that BD diagnosis is a significant predictor of long-term cognitive dysfunction increase (Lewandowski et al., 2011; Torrent et al., 2012). Although there are limited data on the cognitive profile of LOBD (Carlino et al., 2013; Grande et al., 2014), cognitive deficits affecting memory, attention and executive function have been reported for BD patients (Robinson et al., 2006; Osher et al., 2011; Arahamian et al., 2014).

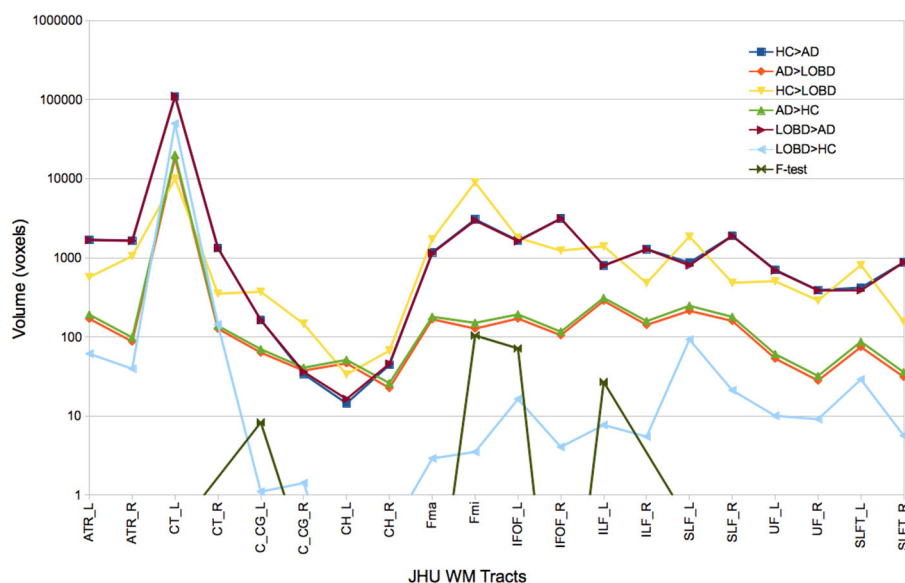


FIGURE 1 | Size of FA skeleton clusters found by each contrast (HC > AD, AD > LOBD, HC > LOBD, AD > HC, LOBD > AD, LOBD > HC, F-test) of the permutation test followed by TFCE cluster inference for each tract identified by the JHU White-Matter Tractography Atlas. R, Right; L, left; ATR, hemispheres of anterior thalamic radiation; CT, corticospinal tract; C.CG, cingulum (cingulate gyrus); CH, Cingulum (hippocampus); Fmi, forceps minor; Fma, forceps major; IFOF, inferior fronto-occipital fasciculus; ILF, inferior longitudinal fasciculus; SLF, superior longitudinal fasciculus; UF, uncinate fasciculus; SLFT, temporal part of SLF.

TBSS Localizations

We focus the discussion on the tracts where the *F*-test finds the greatest clusters for the LOBD > AD contrast, which has quantitatively greater effects than the AD > LOBD contrast. The size of cluster localizations suggest greater axonal degradation in AD than in LOBD in pathways that serve to integrate cognitive and social structures, including tracts that mediate connectivity to frontal and temporal lobes.

In particular, significant differences found in ILF indicate degradation of the fronto-temporal-occipital circuit which is very important for social and emotional processing, leading to behavioral deterioration, which has been assessed as the main discriminant between AD and LOBD (Besga et al., 2015). The ILF tract connects the occipital cortex and temporal lobes including the superior, middle and anterior lobes, mediating the connectivity between three regions: the superior temporal sulcus, the fusiform face area, and the amygdala. Therefore, ILF degradation has impact on the processes of detecting biological motion and eye gaze (Pelphrey and Carter, 2008), as well as facial information processing with social significance, i.e., face identification and facial expression interpretation (Adolphs et al., 1999). We found also significant differences in the cingulate gyrus C.CG, which is part of the cingulate cortex lying above the corpus callosum, and part of the limbic system in charge of processing emotional contents. Together with the previously described impairments of ILF pathway, disruption of C.CG impedes the structural connectivity of an extended circuit that involves frontal, temporal and occipital regions. This circuit controls socio-emotional processing, so its degradation leads to greater behavioral and emotional impairments of AD than LOBD.

The IFOF connects the occipital, posterior temporal, and the orbito-frontal areas (Ashtari, 2012). Simultaneously degraded axonal integrity of left UF, IFOF, and ATR has significant impact on the semantic processing (Han et al., 2013) during specific cognitive tasks related to object recognition. This concurrent effect can explain the increased cognitive impairment of the AD relative to LOBD (Besga et al., 2015), though we have not carried out correlation study of image data with cognitive neuropsychological tests results. The Fmi collects most of the clusters detected, so its relative degradation in AD compared to LOBD is a salient biomarker. The Fmi connects the lateral and medial surfaces of the frontal lobes crossing the midline via the genu of the corpus callosum. Damage of the Fmi detected by decreasing FA in DTI imaging has been associated with fatigue and depression in multiple sclerosis (Gobbi et al., 2014). Degradation of Fmi in mild cognitive impairment (MCI) and AD relative to HC was found in a cohort study using multiple diffusion measures (Alves et al., 2013).

Correlation Analysis

Peripheral biomarkers of inflammation, oxidative stress, and neurotrophins have been related to clinical symptoms, cognitive decline and illness severity in BD (Barbosa et al., 2012; Martinez-Cengotitabengoa et al., 2014) as well as in AD (Berridge, 2013). It has been suggested that inflammation and oxidative stress do not cause AD or LOBD by themselves, but that they reinforce interactions among factors related to these complex neuropsychiatric disorders during brain aging (Forcada et al., 2015), leading to a misbalance between protective and degenerative factors, which predisposes the brain to neurodegenerative diseases (Lewandowski et al., 2011). On the

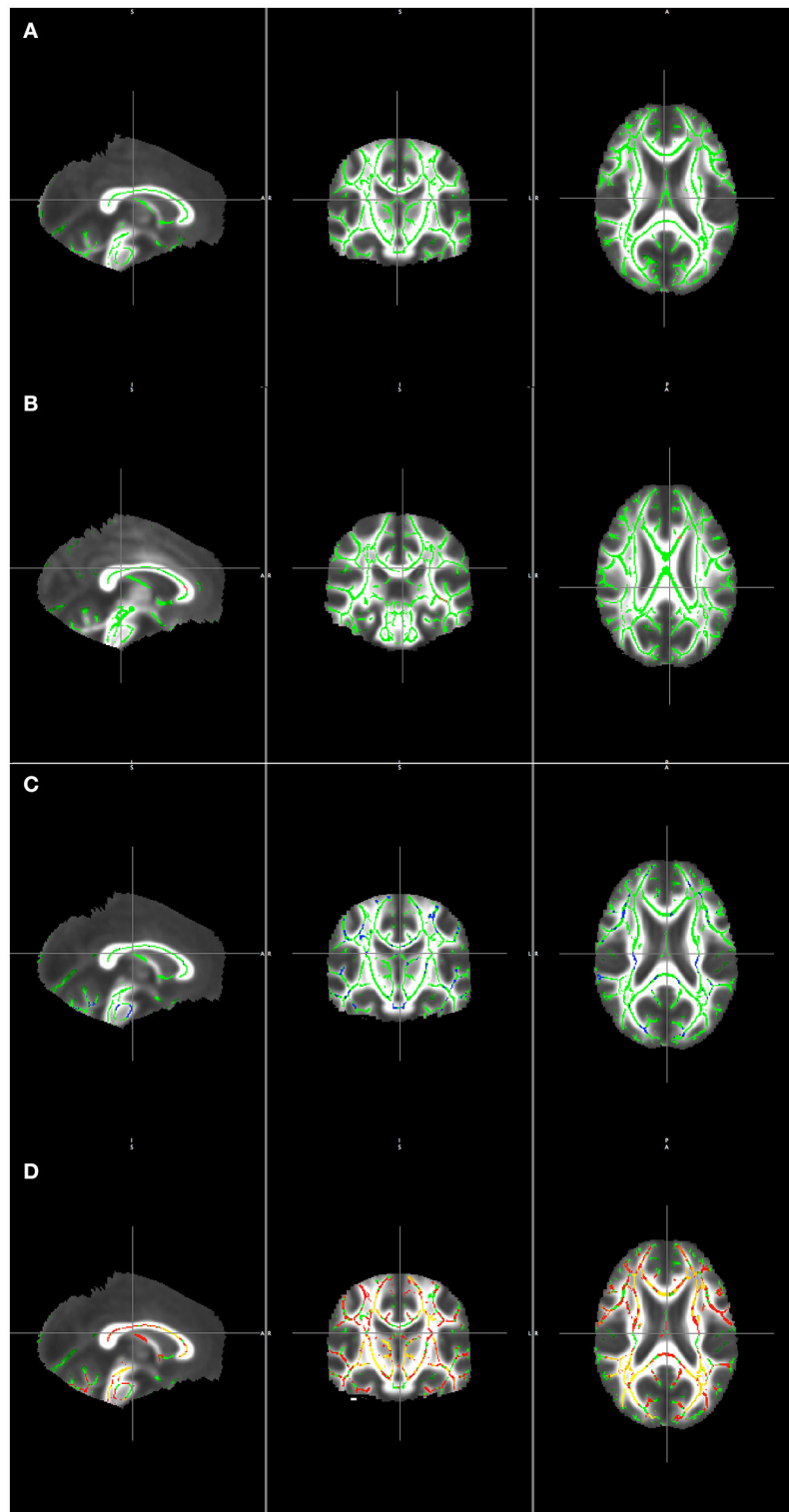


FIGURE 2 | Visualization TBSS detection results masked by the *F*-test overlaid on the mean of registered FA volumes. **(A)** Mean skeleton (green). **(B)** *F* statistics (red-yellow) over the mean skeleton (green). **(C)** Clusters detected from contrast AD > LOBD masked by *F*-test clusters (blue) overlaying the mean skeleton (green). **(D)** Clusters detected from contrast LOBD > AD masked by *F*-test clusters (red) overlaying the mean skeleton (green).

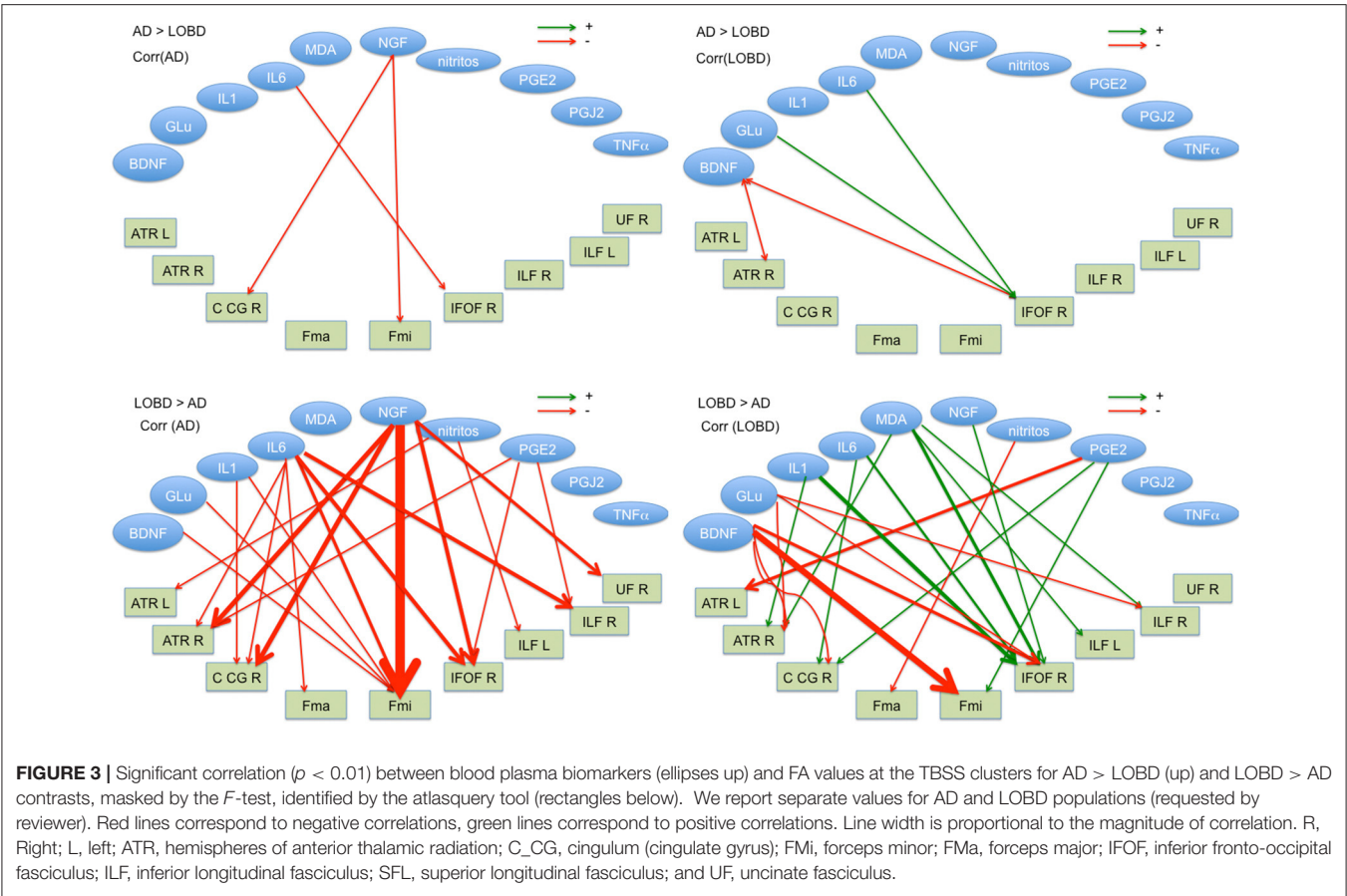


FIGURE 3 | Significant correlation ($p < 0.01$) between blood plasma biomarkers (ellipses up) and FA values at the TBSS clusters for AD > LOBD (up) and LOBD > AD contrasts, masked by the F -test, identified by the atlasquery tool (rectangles below). We report separate values for AD and LOBD populations (requested by reviewer). Red lines correspond to negative correlations, green lines correspond to positive correlations. Line width is proportional to the magnitude of correlation. R, Right; L, left; ATR, hemispheres of anterior thalamic radiation; C_CG, cingulum (cingulate gyrus); Fmi, forceps minor; Fma, forceps major; IFOF, inferior fronto-occipital fasciculus; ILF, inferior longitudinal fasciculus; SFL, superior longitudinal fasciculus; and UF, uncinate fasciculus.

other hand, negative correlation of inflammation biomarkers, i.e., $TNF\alpha$, with FA in the body and isthmus of the corpus callosum has been also found in healthy aging subjects by a TBSS analysis (Arfanakis et al., 2013), showing that systemic inflammation is not necessarily associated with cognitive decline. There are also reports of a significant decrease in BDNF and IL-6 in BD patients at a later stage compared to its early stage, while, inversely, $TNF\alpha$ has a significant increase at the later stage of BD (Kauer-Sant’Anna M, 2009; Grande et al., 2014), suggesting that the inflammation lies in the pathogenesis of BD.

Brain injuries promote the up-regulation of proinflammatory prostaglandins PGE2 (Ahmad et al., 2006), hence blocking the corresponding receptor has been proposed as a target of treatment for stroke and other traumatic brain injuries. Cyclopentenone Prostaglandin PGJ2 is a recently discovered prostaglandin, which has anti-inflammatory functions (Scher and Pillinger, 2005; Zhao et al., 2006), such as the inhibition of a gene in T cells, therefore positive correlation with FA voxels is consistently related to axonal integrity in this area. $TNF\alpha$ is a cytokine involved in systemic inflammation and acute phase reaction, whose role is the regulation of immune cells, inducing inflammation and other effects, such as apoptotic cell death.

The role of NGF as a therapeutic tool for AD has received a lot of attention in the last years (Xu et al., 2016), with strong consideration of the impairment of NGF pathway as cause of AD via the accumulation of amyloid plaques. Clinical trials have been

| | AD > LOBD | | | | LOBD > AD | | | |
|-------------|-----------|---|------|---|-----------|-----|------|----|
| | AD | | LOBD | | AD | | LOBD | |
| | + | − | + | − | + | − | + | − |
| BDNF | | | | 2 | | 3 | | 20 |
| GLU | 1 | | | | | 1 | | 5 |
| IL1 β | | | | | | 5 | | 9 |
| IL6 | | 6 | 1 | | | 44 | | 3 |
| MDA | | | | | | | | 8 |
| NGF | | 1 | | | | 276 | | 1 |
| NO2 | | | | | | 2 | | 2 |
| PGE2 | | | | | | 4 | | 3 |
| PGJ2 | | | | | | | | |
| $TNF\alpha$ | | | | | | | | 1 |

Rows: blood biomarkers. Columns: for each contrast (AD > LOBD, LOBD > AD) and population (AD, LOBD), positive (+) and negative (−) correlation.

carried out^{3,4} studying the effect of NGF gene therapy (Tuszynski et al., 2005). Postmortem analysis showed that the NGF treatment induced response in degenerating neurons exhibited trophic

³<https://clinicaltrials.gov/ct2/show/NCT00017940>
⁴<https://clinicaltrials.gov/ct2/show/NCT01163825>

response without adverse pathological effects (Tuszynski et al., 2015).

In our study, regarding inflammation biomarkers in **Table 1**, we find no effect of the protective PGJ2, and almost no effect of TNF α , while there PGE2 has both positive and negative correlation with LOBD image data, so no conclusion can be given from them. However, though sparsely distributed in different clusters, we find positive correlation of IL1 β and IL6 with LOBD and negative with AD, hence these blood biomarkers are a clear indication of greater inflammation in LOBD pathogenesis.

Regarding oxidative stress biomarkers, we found no differential effect of NO2, because both populations showed the same sizes of negative correlation clusters, but MDA shows positive correlation with LOBD hinting to an added pathogenesis factor. This result is also in agreement with our previous findings using eigenanatomy (Besga et al., 2016). Notice in **Figure 3** that IL1 β , IL6, and MAD affects mostly the IFOF as a cause for behavior impairment.

Regarding neurotrophic factors, we found a big effect of NGF which correlates negatively with AD imaging data, i.e., with the degradation of synaptic integrity in the located tract. We had also a small positive correlation with LOBD that reinforces the value of NGF as a differential diagnostic biomarker between AD and LOBD. This result is in complete agreement with recent AD therapeutic research lines (Tuszynski et al., 2005, 2015; Xu et al., 2016). Notice from **Figure 3** that most of the correlation effects of NGF are located in FMI, suggesting a role in cognitive decline.

CONCLUSIONS

TBSS analysis found widespread white matter disruption in LOBD relative to AD that might be related to axonal integrity degradation measured by decreasing FA in several important tracts. Main effects are located on white matter tracts

that integrate a distributed fronto-temporal-occipital circuit. Disruption of this circuit may be producing the behavioral and cognitive impairments that differentiate LOBD from AD in the clinical and neuropsychological tests. Also, inter-hemispherical tracts FMI has greater axonal integrity degradation in AD than in LOBD, which is a pathophysiological cause for cognitive decline of AD relative to LOBD. Finally, the correlation analysis suggests that neurotrophic factors, i.e., NGF and BDNF, considered together with FA imaging may help to differentiate LOBD from AD. Also, there are indications of greater inflammation (IL1 β , IL6) and oxidative stress (MDA) factors in LOBD than in AD.

AUTHOR CONTRIBUTIONS

AB, IG, AG, and MG have made substantial contributions to the conception or design of the work; AB, DC, MG-L, and JE contributed to the acquisition, analysis, or interpretation of data for the work; all authors contributed drafting the work and revising it critically for important intellectual content; all authors give final approval of the version to be published; all authors agree to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

ACKNOWLEDGMENTS

This research has been partially funded by the Basque Government grant IT874-13 for the research group. DC has been supported by Basque Government post-doctoral grant No. Ref.: POS-2014-1-2, MOD:POSDOC. This work was supported by health research grant 2013111162 from the Department of Education, Linguistic Policy and Culture of the Basque Country Government.

REFERENCES

- Adolphs, R., Tranel, D., Hamann, S., Young, A., and Calder, A. E. (1999). Recognition of facial emotion in nine individuals with bilateral amygdala damage. *Neuropsychologia* 37, 1111–1117. doi: 10.1016/S0028-3932(99)00039-1
- Ahmad, A., Saleem, S., Ahmad, M., and Dore, S. (2006). Prostaglandin ep1 receptor contributes to excitotoxicity and focal ischemic brain damage. *Toxicol. Sci.* 89, 265–270. doi: 10.1093/toxsci/kfj022
- Akiyama, H., Barger, S., Barnum, S., Bradt, B., Bauer, J., Cole, G., et al. (2000). Inflammation and Alzheimer's disease. *Neurobiol. Aging* 21, 383–421. doi: 10.1016/S0197-4580(00)00124-X
- Albert, M., and Blacker, D. (2006). Mild cognitive impairment and dementia. *Annu. Rev. Clin. Psychol.* 2, 379–388. doi: 10.1146/annurev.clinpsy.1.102803.144039
- Alves, G. S., O'Dwyer, L., Jurcoane, A., Oertel-Knochel, V., Knochel, C., Prvulovic, D., et al. (2013). Different patterns of white matter degeneration using multiple diffusion indices and volumetric data in mild cognitive impairment and Alzheimer patients. *PLoS ONE* 7:e52859. doi: 10.1371/journal.pone.0052859
- Andersson, J., Jenkinson, M., and Smith, S. (2007a). *Non-linear Optimisation*. Technical Report TR07JA1, FMRIB.
- Andersson, J., Jenkinson, M., and Smith, S. (2007b). *Non-linear Registration, aka Spatial Normalisation*. Technical Report TR07JA2, FMRIB.
- Aprahamian, I., Ladeira, R., Diniz, B., Forlenza, O., and Nunes, P. (2014). Cognitive impairment in euthymic older adults with bipolar disorder: a controlled study using cognitive screening tests. *Am. J. Geriatr. Psychiatry* 22, 389–397. doi: 10.1016/j.jagp.2012.08.013
- Arfanakis, K., Fleischman, D. A., Grisot, G., Barth, C. M., Varentsova, A., Morris, M. C., et al. (2013). Systemic inflammation in non-demented elderly human subjects: brain microstructure and cognition. *PLoS ONE* 8:e73107. doi: 10.1371/journal.pone.0073107
- Ashtari, M. (2012). Anatomy and functional role of the inferior longitudinal fasciculus: a search that has just begun. *Develop. Med. Child Neurol* 54, 6–7. doi: 10.1111/j.1469-8749.2011.04122.x
- Avants, B., Dhillon, P., Kandel, B., Cook, P., McMillan, C., Grossman, M., et al. (2012). Eigenanatomy improves detection power for longitudinal cortical change. *Med. Image Comput. Comput. Assist. Interv.* 15, 206–213.
- Avants, B. B., Libon, D. J., Rascovsky, K., Boller, A., McMillan, C. T., Massimo, L., et al. (2014). Sparse canonical correlation analysis relates network-level atrophy to multivariate cognitive measures in a neurodegenerative population. *Neuroimage* 84, 698–711. doi: 10.1016/j.neuroimage.2013.09.048
- Bach, M., Laun, F. B., Leemans, A., Tax, C. M., Biessels, G. J., Stieltjes, B., et al. (2014). Methodological considerations on tract-based spatial statistics (tbss). *Neuroimage* 100, 358–369. doi: 10.1016/j.neuroimage.2014.06.021
- Barbosa, I., Rocha, N., Huguet, R., Ferreira, R., Salgado, J., and Carvalho, L. A., (2012). Executive dysfunction in euthymic bipolar disorder patients

- and its association with plasma biomarkers. *J. Affect. Disord.* 137, 151–155. doi: 10.1016/j.jad.2011.12.034
- Basser, P. J., Mattiello, J., and LeBihan, D. (1994). MR diffusion tensor spectroscopy and imaging. *Biophys. J.* 66, 259–267. doi: 10.1016/S0006-3495(94)80775-1
- Bauer, I. E., Pascoe, M., Wollenhaupt-Aguiar, B., Kapczynski, F., and Soares, J. C. (2014). Inflammatory mediators of cognitive impairment in bipolar disorder. *J. Psychiat. Res.* 56, 18–27. doi: 10.1016/j.jpsychires.2014.04.017
- Bauer, M., Glenn, T., Alda, M., Andreassen, O., Angelopoulos, E., Arda, R., et al. (2015a). Influence of birth cohort on age of onset cluster analysis in bipolar i disorder. *Euro. Psychiatry* 30, 99–105. doi: 10.1016/j.eurpsy.2014.10.005
- Bauer, M., Glenn, T., Alda, M., Andreassen, O. A., Angelopoulos, E., Arda, R., et al. (2015b). Influence of light exposure during early life on the age of onset of bipolar disorder. *J. Psychiat. Res.* 64, 1–8. doi: 10.1016/j.jpsychires.2015.03.013
- Bauer, M., Glenn, T., Alda, M., Andreassen, O. A., Angelopoulos, E., Arda, R., et al. (2014). Relationship between sunlight and the age of onset of bipolar disorder: an international multisite study. *J. Affect. Disord.* 167, 104–111. doi: 10.1016/j.jad.2014.05.032
- Behrens, T. E., Woolrich, M. W., Jenkinson, M., Johansen-Berg, H., Nunes, R. G., Clare, S., et al. (2003). Characterization and propagation of uncertainty in diffusion-weighted MR imaging. *Magn. Reson. Med.* 50, 1077–1088. doi: 10.1002/mrm.10609
- Berlyand, Y., Weintraub, D., Xie, S. X., Mellis, I. A., Doshi, J., Rick, J., et al. (2016). An Alzheimer's disease-derived biomarker signature identifies parkinson's disease patients with dementia. *PLoS ONE* 11:e0147319. doi: 10.1371/journal.pone.0147319
- Besga, A., Chyzhyk, D., Gonzalez-Ortega, I., Savio, A., Ayerdi, B., Echeveste, J., et al. (2016). Eigenanatomy on fractional anisotropy imaging provides white matter anatomical features discriminating between Alzheimer's disease and late onset bipolar disorder. *Curr. Alzheimer Res.* 13, 557–565. doi: 10.2174/1567205013666151116125349
- Besga, A., Gonzalez-Ortega, I., Echeburua, E., Savio, A., Ayerdi, B., Chyzhyk, D., et al. (2015). Discrimination between Alzheimer's disease and late onset bipolar disorder using multivariate analysis. *Front. Aging Neurosci.* 7:231. doi: 10.3389/fnagi.2015.00231
- Besga, A., Martinez-Cengotitabengoa, M., Gonzalez-Ortega, I., Gutierrez, M., Barbeito, S., and Gonzalez-Pinto, A. (2011). The role of white matter damage in late onset bipolar disorder. *Maturitas* 70, 160–163. doi: 10.1016/j.maturitas.2011.07.005
- Besga, A., Termenon, M., Graña, M., Echeveste, J., Perez, J., and Gonzalez-Pinto, A. (2012). Discovering Alzheimer's disease and bipolar disorder white matter effects building computer aided diagnostic systems on brain diffusion tensor imaging features. *Neurosci. Lett.* 520, 71–76. doi: 10.1016/j.neulet.2012.05.033
- Brydon, L., Walker, C., Wawrzyniak, A., Whitehead, D., Okamura, H., Yajima, J., et al. (2009). Synergistic effects of psychological and immune stressors on inflammatory cytokine and sickness responses in humans. *Brain Behav. Immun.* 23, 217–224. doi: 10.1016/j.bbi.2008.09.007
- Carlino, A., Stinnett, J., and Kim, D. (2013). New onset of bipolar disorder in late life. *Psychosomatics* 54, 94–97. doi: 10.1016/j.psych.2012.01.006
- Castanon, N., Lasselin, J., and Capuron, L. (2014). Neuropsychiatric comorbidity in obesity: role of inflammatory processes. *Front. Endocrinol. (Lausanne)* 5:74. doi: 10.3389/fendo.2014.00074
- Chou, P.-H., Tseng, W.-J., Chen, L.-M., Lin, C.-C., Lan, T.-H., and Chan, C.-H. (2015). Late onset bipolar disorder: a case report and review of the literature. *J. Clin. Gerontol. Geriatr.* 6, 27–29. doi: 10.1016/j.jcgg.2014.05.002
- Depp, C., and Jeste, D. (2004). Bipolar disorder in older adults: a critical review. *Bipolar Disord.* 6, 343–367. doi: 10.1111/j.1399-5618.2004.00139.x
- Dickerson, F., Stallings, C., Origoni, A., Vaughan, C., Katsafanas, E., Khushalani, S., et al. (2013). A combined marker of inflammation in individuals with mania. *PLoS ONE* 8:e73520. doi: 10.1371/journal.pone.0073520
- Forcada, I., Mur, M., Mora, E., Vieta, E., Bartras-Faz, D., and Portella, M. J. (2015). The influence of cognitive reserve on psychosocial and neuropsychological functioning in bipolar disorder. *Euro. Neuropsychopharmacol.* 25, 214–222. doi: 10.1016/j.euroneuro.2014.07.018
- Garcia-Bueno, M., Bioque, M., Mac-Dowell, K. S., Barcones, M. F., Martínez-Cengotitabengoa, M., Pina-Camacho, L., et al. (2014). Proanti-inflammatory dysregulation in patients with first episode of psychosis: toward an integrative inflammatory hypothesis of schizophrenia. *Schizophr. Bull.* 40, 376–387. doi: 10.1093/schbul/sbt001
- Gobbi, C., Rocca, M., Pagani, E., Riccitelli, G., Pravata, E., Radaelli, M., et al. (2014). Forceps minor damage and co-occurrence of depression and fatigue in multiple sclerosis. *Multiple Scleros.* J. 20, 1633–1640. doi: 10.1177/1352458514530022
- Goldstein, B., Kemp, D., Soczynska, J., and McIntyre, R. (2009). Inflammation and the phenomenology, pathophysiology, comorbidity, and treatment of bipolar disorder: a systematic review of the literature. *J. Clin. Psychiatry* 70, 1078–1090. doi: 10.4088/JCP.08r04505
- Graña, M., Termenon, M., Savio, A., Gonzalez-Pinto, A., Echeveste, J., Perez, J., et al. (2011). Computer aided diagnosis system for Alzheimer disease using brain diffusion tensor imaging features selected by pearson's correlation. *Neurosci. Lett.* 502, 225–229. doi: 10.1016/j.neulet.2011.07.049
- Grande, I., Magalhaes, P., Chendo, I., Stertz, L., Panizutti, B., Colpo, G., et al. (2014). Staging bipolar disorder: clinical, biochemical, and functional correlates. *Acta Psychiatr. Scand.* 129, 437–444. doi: 10.1111/acps.12268
- Han, Z., Ma, Y., Gong, G., He, Y., Caramazza, A., and Bi, Y. (2013). White matter structural connectivity underlying semantic processing: evidence from brain damaged patients. *Brain* 136, 2952–2965. doi: 10.1093/brain/awt205
- Hope, S., Hoseth, E., Dieset, I., Mørch, R. H., Aas, M., Aukrust, P., et al. (2015). Inflammatory markers are associated with general cognitive abilities in schizophrenia and bipolar disorder patients and healthy controls. *Schizophr. Res.* 165, 188–194. doi: 10.1016/j.schres.2015.04.004
- Kamer, A., Craig, R., Dasanayake, A., Brys, M., Glodzik-Sobanska, L., and de Leon, M. (2008). Inflammation and Alzheimer's disease: possible role of periodontal diseases. *Alzheimer Dement.* 4, 242–250. doi: 10.1016/j.jalz.2007.08.004
- Kauer-Sant'Anna, M., Kapczynski, F., Andreazza, A. C., Bond, D. J., Lam, R. W., Young, L. T., et al. (2009). Brain-derived neurotrophic factor and inflammatory markers in patients with early- vs. late-stage bipolar disorder. *Int. J. Neuropsychopharmacol.* 12, 447–458. doi: 10.1017/S1461145708009310
- Kawas, C., Corrada, M., Brookmeyer, R., Morrison, A., Resnick, S., Zonderman, A., et al. (2003). Visual memory predicts Alzheimer's disease more than a decade before diagnosis. *Neurology* 60, 1089–1093. doi: 10.1212/01.WNL.0000055813.36504.BF
- Kendall, M. G., and Stuart, A. (1973). *The Advanced Theory of Statistics, Volume 2: Inference and Relationship*. London: Griffin.
- Konradi, C., Sullivan, S., and Clay, H. (2012). Mitochondria, oligodendrocytes and inflammation in bipolar disorder: evidence from transcriptome studies points to intriguing parallels with multiple sclerosis. *Neurobiol. Dis.* 45, 37–47. doi: 10.1016/j.nbd.2011.01.025
- Lebert, F., Lys, H., Haem, E., and Pasquier, F. (2008). Dementia following bipolar disorder. *Encephale* 34, 606–610. doi: 10.1016/j.encep.2007.12.007
- Leboyer, M., Soreca, I., Scott, J., Frye, M., Henry, C., Tamouza, R., et al. (2012). Can bipolar disorder be viewed as a multi-system inflammatory disease? *J. Affect. Disord.* 141, 1–10. doi: 10.1016/j.jad.2011.12.049
- Lee, S.-Y., Chen, S.-L., Chang, Y.-H., Chen, P., and Huang, S.-Y. (2013). Inflammation's association with metabolic profiles before and after a twelve-week clinical trial in drug-naïve patients with bipolar ii disorder. *PLoS ONE* 8:e66847. doi: 10.1371/journal.pone.0066847
- Lewandowski, K., Cohen, B., and Ongur, D. (2011). Evolution of neuropsychological dysfunction during the course of schizophrenia and bipolar disorder. *Psychol. Med.* 41, 225–241. doi: 10.1017/S0033291710001042
- Martinez-Cengotitabengoa, M., Mico, J., Arango, C., Castro-Fornieles, J., Graell, M., Paya, B., et al. (2014). Basal low antioxidant capacity correlates with cognitive deficits in early onset psychosis: a 2-year follow-up study. *Schizophr. Res.* 156, 23–29. doi: 10.1016/j.schres.2014.03.025
- Tuszynski, M. H., Yang, J. H., Barba, D., U. H. S., Bakay, R. A., Pay, M. M., et al. (2015). Nerve growth factor gene therapy: activation of neuronal responses in Alzheimer disease. *JAMA Neurol.* 72, 1139–1147. doi: 10.1001/jamaneurol.2015.1807
- Berridge, M. J. (2013). Dysregulation of neural calcium signaling in Alzheimer disease, bipolar disorder and schizophrenia. *Prion* 7, 2–13. doi: 10.4161/pri.21767
- Osher, Y., Dobron, A., Belmaker, R., Bersudsky, Y., and Dwolatzky, T. (2011). Computerized testing of neurocognitive function in euthymic bipolar patients compared to those with mild cognitive impairment and cognitively healthy controls. *Psychosomat.* 80, 298–303. doi: 10.1159/000324508
- Pearson, K. (1895). Notes on regression and inheritance in the case of two parents. *Proc. R. Soc. Lond.* 58, 240–242. doi: 10.1098/rspl.1895.0041

- Pelphrey, K., and Carter, E. (2008). Charting the typical and atypical development of the social brain. *Dev. Psychopathol.* 20, 1081–1102. doi: 10.1017/S0954579408000515
- Pierpaoli, C., Jezzard, P., Basser, P. J., Barnett, A., and Chiro, G. D. (1996). Diffusion tensor MR imaging of the human brain. *Radiology* 201, 637–648. doi: 10.1148/radiology.201.3.8939209
- Prabhakar, D., and Balon, R. (2010). Late-onset bipolar disorder a case for careful appraisal. *Psychiatry (Edgmont)* 7, 34–37.
- Ramirez, J., Gorriz, J., Segovia, F., Chaves, R., Salas-Gonzalez, D., Lopez, M., et al. (2010). "Early Alzheimer's disease diagnosis using partial least squares and random forests," in *Biomedical Imaging: From Nano to Macro, 2010 IEEE International Symposium on* (Rotterdam), 81–84.
- Robinson, L., Thompson, J., Gallagher, P., Goswami, U., Young, A., Ferrier, I. N., et al. (2006). A meta-analysis of cognitive deficits in euthymic patients with bipolar disorder. *J. Affect. Disord.* 93, 105–115. doi: 10.1016/j.jad.2006.02.016
- Salas-Gonzalez, D., Górriz, J. M., Ramírez, J., López, M., Illan, I. A., Segovia, F., et al. (2009). Analysis of SPECT brain images for the diagnosis of Alzheimer's disease using moments and support vector machines. *Neurosci. Lett.* 461, 60–64. doi: 10.1016/j.neulet.2009.05.056
- Sardi, F., Fassina, L., Venturini, L., Inguscio, M., Guerriero, F., Rolfo, E., et al. (2011). Alzheimer's disease, autoimmunity and inflammation. the good, the bad and the ugly. *Autoimmun. Rev.* 11, 149–153. doi: 10.1016/j.autrev.2011.09.005
- Savio, A., Garcia-Sebastian, M., Chyzyk, D., Hernandez, C., Graña, M., Sistiaga, A., et al. (2011). Neurocognitive disorder detection based on feature vectors extracted from vbm analysis of structural mri. *Comput. Biol. Med.* 41, 600–610. doi: 10.1016/j.compbiomed.2011.05.010
- Scher, J. U., and Pillinger, M. H. (2005). 15d-pgij2: The anti-inflammatory prostaglandin? *Clin. Immunol.* 114, 100–109. doi: 10.1016/j.clim.2004.09.008
- Sigut, J., Piñeiro, J., Gonzalez, E., and Torres, J. (2007). An expert system for supervised classifier design: application to Alzheimer diagnosis. *Expert Syst. Appl.* 32, 927–938. doi: 10.1016/j.eswa.2006.01.026
- Singhal, G., Jaehne, E., Corrigan, F., Toben, C., and Baune, B. (2014). Inflammasomes in neuroinflammation and changes in brain function: a focused review. *Front. Neurosci.* 8:315. doi: 10.3389/fnins.2014.00315
- Smith, S. M., Jenkinson, M., Johansen-Berg, H., Rueckert, D., Nichols, T. E., Mackay, C. E., et al. (2006). Tract-based spatial statistics: Voxelwise analysis of multi-subject diffusion data. *Neuroimage* 31, 1487–1505. doi: 10.1016/j.neuroimage.2006.02.024
- Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E., Johansen-Berg, H., et al. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage*, 23(Suppl. 1), S208–S219. Mathematics in Brain Imaging. doi: 10.1016/j.neuroimage.2004.07.051
- Smith, S. M., and Nichols, T. E. (2009). Threshold-free cluster enhancement: addressing problems of smoothing, threshold dependence and localisation in cluster inference. *Neuroimage* 44, 83–98. doi: 10.1016/j.neuroimage.2008.03.061
- Sparding, T., Silander, K., Palsson, E., Ostlind, J., Sellgren, C., Ekman, C. J., et al. (2015). Cognitive functioning in clinically stable patients with bipolar disorder i and ii. *PLoS ONE* 10:e0115562. doi: 10.1371/journal.pone.0115562
- Termenon, M., Graña, M., Besga, A., Echeveste, J., and Gonzalez-Pinto, A. (2013). Lattice independent component analysis feature selection on diffusion weighted imaging for Alzheimer's disease classification. *Neurocomputing* 114, 132–141. doi: 10.1016/j.neucom.2012.08.044
- Torrent, C., Martinez-Aran, A., del Mar Bonnin, C., Reinares, M., Daban, C., Sole, B., et al. (2012). Long-term outcome of cognitive impairment in bipolar disorder. *J. Clin. Psychiatry* 73, e899–e905. doi: 10.4088/JCP.11m07471
- Tuszynski, M. H., Thal, L., Pay, M., Salmon, D. P., U. H. S., Bakay, R., et al. (2005). A phase 1 clinical trial of nerve growth factor gene therapy for Alzheimer disease. *Nat. Med.* 11, 551–555. doi: 10.1038/nm1239
- Weintraub, S., Wicklund, A., and Salmon, D. (2012). The neuropsychological profile of Alzheimer disease. *Cold Spring Harb. Perspect. Med.* 2:a006171. doi: 10.1101/cshperspect.a006171
- Westman, E., Simmons, A., Zhang, Y., Muehlboeck, J., Tunnard, C., Liu, Y., et al. (2011). Multivariate analysis of mri data for Alzheimer's disease, mild cognitive impairment and healthy controls. *Neuroimage* 54, 1178–1187. doi: 10.1016/j.neuroimage.2010.08.044
- Xu, C.-J., Wang, J.-L., and Jin, W.-L. (2016). The emerging therapeutic role of ngf in Alzheimer's disease. *Neurochem. Res.* 41, 1211–1218. doi: 10.1007/s11064-016-1829-9
- Zahodne, L. B., Ornstein, K., Cosentino, S., Devanand, D., and Stern, Y. (2015). Longitudinal relationships between Alzheimer disease progression and psychosis, depressed mood, and agitation/aggression. *Am. J. Geriatr. Psychiatry* 23, 130–140. doi: 10.1016/j.jagp.2013.03.014
- Zanetti, M. V., Cordeiro, Q., and Busatto, G. F. (2007). Late onset bipolar disorder associated with white matter hyperintensities: a pathophysiological hypothesis. *Progr. Neuro Psychopharmacol. Biol. Psychiatry* 31, 551–556.
- Zhao, X., Zhang, Y., Strong, R., Grotta, J. C., and Aronowski, J. (2006). 15d-prostaglandin j2 activates peroxisome proliferator-activated receptor- γ , promotes expression of catalase, and reduces inflammation, behavioral dysfunction, and neuronal loss after intracerebral hemorrhage in rats. *J. Cereb. Blood Flow Metab.* 26, 811–820. doi: 10.1038/sj.jcbfm.9600233

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.

Copyright © 2017 Besga, Chyzyk, Gonzalez-Ortega, Echeveste, Graña-Lecuona, Graña and Gonzalez-Pinto. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Retrospective Diagnosis of Parkinsonian Syndromes Using Whole-Brain Atrophy Rates

Carlos Guevara*, Kateryna Bulatova, Wendy Soruco, Guido Gonzalez and Gonzalo A. Farías

Facultad de Medicina, Universidad de Chile, Santiago, Chile

Objective: The absence of markers for *ante-mortem* diagnosis of idiopathic Parkinson's disease (IPD), multiple system atrophy (MSA), and progressive supranuclear palsy (PSP) results in these disorders being commonly mistaken for each other, particularly in the initial stages. We aimed to investigate annualized whole-brain atrophy rates (a-WBAR) in these disorders to aid in the diagnosis between IPD vs. PSP and MSA.

Methods: Ten healthy controls, 20 IPD, 39 PSP, and 41 MSA patients were studied using Structural Imaging Evaluation with Normalization of Atrophy (SIENA). SIENA is an MRI-based algorithm that quantifies brain tissue volume and does not require radiotracers. SIENA has been shown to have a low estimation error for atrophy rate over the whole brain (0.5%).

Results: In controls, the a-WBAR was $0.37\% \pm 0.28$ (CI 95% 0.17–0.57), while in IPD a-WBAR was $0.54\% \pm 0.38$ (CI 95% 0.32–0.68). The IPD patients did not differ from the controls. In PSP, the a-WBAR was $1.93\% \pm 1.1$ (CI 95% 1.5–2.2). In MSA a-WBAR was $1.65\% \pm 0.9$ (CI 95% 1.37–1.93). MSA did not differ from PSP. The a-WBAR in PSP and MSA were significantly higher than in IPD ($p < 0.001$). a-WBAR 0.6% differentiated patients with IPD from those with PSA and MSA with 91% sensitivity and 80% specificity.

Conclusions: a-WBAR within the normal range is unlikely to be observed in PSP or MSA. a-WBAR may add a potential retrospective application to improve the diagnostic accuracy of MSA and PSP vs. IPD during the first year of clinical assessment.

Keywords: whole brain atrophy rate, multiple system atrophy, progressive supranuclear palsy, idiopathic Parkinson's disease

OPEN ACCESS

Edited by:

Javier Ramírez,
University of Granada, Spain

Reviewed by:

Johannes Levin,
Ludwig-Maximilians-Universität
München, Germany
Alessia Sarica,
Institute of Molecular Bioimaging and
Physiology (CNR), Italy

*Correspondence:

Carlos Guevara
neurocrs@hotmail.com

Received: 05 February 2017

Accepted: 29 March 2017

Published: 19 April 2017

Citation:

Guevara C, Bulatova K, Soruco W,
Gonzalez G and Farías GA (2017)
Retrospective Diagnosis of
Parkinsonian Syndromes Using
Whole-Brain Atrophy Rates.
Front. Aging Neurosci. 9:99.
doi: 10.3389/fnagi.2017.00099

INTRODUCTION

Multiple system atrophy (MSA) and progressive supranuclear palsy (PSP)—sometimes designated as “Parkinson plus syndromes”—are debilitating neurodegenerative disorders with heterogeneous presentation, inexorable progression, and a median survival of between 5 and 10 years. There is a need to improve the differentiation between idiopathic Parkinson's disease (IPD) and MSA vs. PSP. PSP and MSA can be misdiagnosed as IPD (and vice versa), especially in early stages, as these disorders share some common clinical features, such as bradykinesia and rigidity and even initial response to levodopa treatment, making the diagnosis, which is initially based on clinical presentation only, rather uncertain. Indeed, in 2004 Adler et al. reported that only 26% of IPD cases with signs and

symptoms present for <5 years had neuropathologic confirmation (Adler et al., 2014). Although a number of neuroimaging techniques allow for partial distinction among these diseases (Politis, 2014), no neuroimaging modalities are specifically recommended for routine use in clinical practice for the differential diagnosis between IPD vs. MSA and PSP.

Whole brain atrophy rates (WBAR) from magnetic resonance imaging (MRI) data may be an informative way to quantify disease progression in an unbiased fashion. This approach reduces inter-individual variability in brain size and morphology when baseline scans are used as reference point so that the subject acts as his or her own control. This avenue has been extensively explored in Alzheimer disease, (Fox and Freeborough, 1997; Schott et al., 2005; Ridha et al., 2008; Sluimer et al., 2008a,b) and also in other degenerative dementias such as frontotemporal dementia (Chan et al., 2001; Gordon et al., 2010) and Huntington disease (Hobbs et al., 2010). For normal aging, the annualized-WBAR (a-WBAR) has been estimated to be below 0.6% (Josephs et al., 2006; Whitwell et al., 2007; Sluimer et al., 2008a). To date few studies have used such techniques in PSP and MSA with small numbers of patients. In six autopsy-confirmed PSP cases (Josephs et al., 2006), the a-WBAR [measured using the boundary shift integral (BSI; Freeborough et al., 1997), a (semi-) automated technique] was 1.3%. In another five proven PSP cases, this figure was 1% (Josephs et al., 2006; Whitwell et al., 2007); in another study, also using BSI, a-WBAR estimates were approximately 1% for both PSP and MSA based on 17 PSP cases and 9 cases with MSA-P (Paviour et al., 2006).

An alternative method is provided by structural image evaluation, using normalization, of atrophy (SIENA; Smith et al., 2001, 2002; <http://www.fmrib.ox.ac.uk/analysis/research/siena>). The suitability of SIENA for longitudinal studies is based on: (a) it is direct, based on a registration of two scans taken at different time points, without the confounding effects of choice of a “template” to which to register, (b) all the stages are fully automated; (c) it has been shown to be robust to changes in acquisition parameters including pulse sequence and slice thickness (Smith et al., 2001), which is an important advantage in clinical trials which are usually multi-center. SIENA has been shown to have a low estimation error for atrophy rate over the whole brain (0.5%; Smith et al., 2001, 2002).

In this study, we used SIENA to estimate a-WBAR in IPD, PSP, and MSA. We aimed to explore the retrospective application of a-WBAR to differentiate IPD from MSA and PSP, after 1 year from the baseline assessment and before 5 years of the disease course.

MATERIALS AND METHODS

Subjects and Clinical Assessment

One hundred and ten participants (10 healthy controls, 20 IPD without dementia, 39 PSP, and 41 MSA patients) were recruited from the Movement Disorders Clinic at the Hospital San Juan de Dios, Santiago, Chile. Internationally established operational criteria were used to assess the diagnoses of MSA, PSP, and IPD (Wenning et al., 1997; Hughes et al., 2001; Litvan et al., 2003). Controls were independently functioning community dwellers, did not have active neurologic or psychiatric conditions, did

not have cognitive complaints, and had a normal neurological examination. Fourteen IPD patients had the tremor dominant phenotype and six had the postural instability gait disorder phenotype. Of the 39 PSP patients, 30 had the typical features of classic PSP (Richardson's syndrome). Nine patients were clinically classified as having atypical profiles: four with tremor and moderate L-dopa responsiveness (PSP-Parkinsonism variant), three PSP with corticobasal syndrome (PSP-CBS) and two PSP with progressive nonfluent aphasia (PSP-PNFA; Respondek et al., 2014). Thirty-five probable MSA patients were categorized as MSA-P (predominant Parkinsonian features) and six as MSA-C (predominant cerebellar features). All participants were assessed on their usual dopaminergic medication and the IPD patients were evaluated in the “on state.” The patients' demographics and clinical variables are presented in **Table 1**.

Clinical parameters were explored using the 18-item Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS) motor symptoms (UPDRS III; Goetz et al., 2007) and the Hoehn and Yahr Scale (H&Y; Hoehn and Yahr, 1967), and executive function was assessed using the Frontal Assessment Battery (FAB; Dubois et al., 2000).

MRI Acquisition

Between 2012 and 2015, patients underwent an MRI brain scan. MRI images were acquired on a 3.0 T Philips Medical System. Axial T₁-weighted images of the whole brain were obtained using a 3D inversion recovery prepared spoiled gradient echo (IR-SPGR) sequence. The following parameters were used: repetition time of 8.1 ms; echo time of 3.7 ms; inversion time of 450 ms; voxel size of 0.699 × 0.699 × 1 mm; excitation flip angle of 8°; matrix size of 248 × 226; field of view of 24 cm; and 198 axial slice of 1 mm. An experienced neuroradiologist (GG) assessed the MRI scans of every patient to rule out gross anatomical abnormalities. Patients underwent a second MRI brain scan at the time of the last study visit (12 months after the baseline scan). Subjects were included in the study if they had two MRI scans of adequate quality and the brain extraction step in SIENA functioned correctly. None of the MRI images included in this study showed any structural abnormalities other than atrophy-related changes. These inclusion criteria were assessed by a visual inspection of the raw and processed data for each patient scan. For both the baseline and follow-up assessments, the clinical data and MRI scans were acquired within 1 week of each other. The mean scan interval was 1.04 ± 0.07 years.

Data Analysis

All of the images were converted in NIFTI format using MRIcron software (<http://people.cas.sc.edu/rorden/mricron/dcm2nii.html>) in preparation for processing using SIENA. Before further processing, all of the data were anonymized by removing any reference to the patients' names from the image headers and ensuring that the file names were based on a unique ID rather than any of the patients' personal details, including their clinical group. The SIENA processing algorithm has been validated and described in detail elsewhere (Smith et al., 2002). Briefly, the processing stages are as follows: (1) Brain extraction

TABLE 1 | Baseline demographics, clinical features, and a-WBAR.

| | Controls N = 10 | IPD N = 20 | PSP N = 39 | MSA N = 41 | Group comparisons | Significant pair-wise comparison |
|--|-----------------------------|-----------------------------|-----------------------------|---------------------------|--|--|
| Age (years) ^a Mean ± SD | 64.6 ± 9.9 | 62.2 ± 11.5 | 68.2 ± 6.3 | 60.4 ± 7.7 | $F = 6.18$ $df = 3$ $p = 0.01$ | PSP vs. MSA < 0.001 |
| Gender (M:F) ^b | 3:7 | 8:12 | 21:18 | 32:9 | $\chi^2 = 13$ $df = 3$ $p = 0.005$ | |
| Disease duration ^a (years) Mean ± SD | N/A | 3.1 ± 3.3 | 3.0 ± 1.7 | 4.3 ± 2.3 | $F = 3.4$ $df = 2$ $p = 0.03$ | PSP vs. MSA = 0.04 |
| a-WBAR ^a (Mean ± SD plus 95% confident interval) | 0.37% ± 0.28 (0.17–0.57) | 0.54% ± 0.38 (0.32–0.68) | 1.93% ± 1.1 (1.5–2.2) | 1.65% ± 0.9 (1.3–1.9) | $F = 16$ $df = 3$ $p < 0.001$ | IPD vs. MSA < 0.001 IPD vs. PSP < 0.001 |
| Baseline UPDRS III ^c (median score plus range) | | 23.2 ± 12 (3–46) | 31 ± 13 (6–62) | 36.1 ± 18 (10–67) | $\chi^2 = 7$ $df = 2$ $p = 0.03$ | IPD vs. MSA = 0.01 |
| Annualized UPDRS change (Mean ± SD plus range) | | 1.18 ± 12.7 (–4.7–7.1) | 5.6 ± 9.5* (2.5–8.7) | 6.1 ± 6.9* (3.9–8.6) | | |
| H & Y ^c (Mean ± SD score plus range) | | 1.9 ± 0.6 (1.0–3.0) | 3.0 ± 0.8 (2.0–5.0) | 3.0 ± 0.9 (1.0–5.0) | $\chi^2 = 27$ $df = 2$ $p < 0.001$ | IPD vs. MSA < 0.001 IPD vs. PSP < 0.001 |
| Annualized H & Y change (Mean ± SD plus range) | | 0.1 ± 0.6 (–0.15–0.4) | 0.45 ± 0.55* (0.27–0.63) | 0.4 ± 0.4* (0.28–0.6) | | |
| FAB ^c | | 14.6 ± 3.5 (5–18) | 8.9 ± 4.2 (0–16) | 14.0 ± 3.3 (4–18) | $\chi^2 = 19$ $df = 2$ $p = 0.016$ | IPD vs. PSP < 0.001 MSA vs. PSP < 0.001 |
| Annualized FAB change (Mean ± SD plus range) | | –0.1 ± 1.9 (–1.0–0.8) | –0.46 ± 3.6* (–1.8–0.9) | –0.6 ± 2.08* (–1–0.17) | | |

^aANOVA test.^bChi square test.^cKruskal-Wallis test and post hoc procedure with MannWhitney test $p = 0.05/3 = 0.016$.* Difference between baseline and repeat score with a $p < 0.05$ (Wilcoxon's signed rank test).

UPDRS III, Unified Parkinson's Disease Rating Scale Part III; H&Y, Hoehn & Yahr Scale; FAB, Frontal Assessment Battery; a-WBAR, annual whole-brain atrophy rate.

(BET): segmentation of the brain from non-brain tissue for each scan, followed by skull extraction. (2) Registration: the segmented brain from the second (follow-up) scan is registered to that of the first (baseline) using a linear transformation. The two skull images are used as normalizing factors to constrain the scale and skew. (3) Tissue type segmentation: white matter and gray matter tissues are treated as one tissue and the cerebrospinal fluid as another. (4) Change analysis: detection of the brain edges on both registered brain images and then estimation of the motion of the brain surface edges. The direction of movement from the first image to second image indicates whether atrophy or growth has occurred. Finally, the percentage of global brain volume change is obtained for each subject from the mean of all of the edge point motions.

Statistical Analyses

Statistical analyses of the clinical data and clinical-imaging correlations were performed using the Statistical Package for

Social Sciences (SPSS, Inc., Chicago, IL, USA, version 22). The results are presented as the mean ± SD. In all cases, a two-sided $p < 0.05$ was considered significant. Visual inspection of the data using histograms and QQ-plots was performed to test for violations of the assumption of a normal distribution. Levene's test of equal variances was used to verify the assumption of the homogeneity of variances. Because of these verifications, parametric and non-parametric statistical tests were used. One-way analysis of variance was performed for normally distributed data (age at examination, disease duration, a-WBAR). The Tukey test was used to control for multiple testing. Because disease severity and neuropsychological measures were non-normally distributed, between group differences were compared using Kruskal-Wallis tests, and when necessary, a *post hoc* procedure with Bonferroni correction for multiple tests ($p = 0.05$ was divided by 3) was used to compare the four disease groups. A χ^2 -test for homogeneity was used to compare the distribution of males and females across groups. The a-WBAR was calculated by

dividing the WBAR values by the interscan interval in years. Clinical scores were also annualized by dividing the unit change between the assessments by years. Difference between baseline and repeat score were assessed using the Wilcoxon's signed rank test.

A-WBAR cut-off points for the differentiation between groups were determined by the Receiver Operating Characteristic curve (ROC) to define maximal sum of sensitivity and specificity.

Standard Protocol Approval, Registrations, and Patient Consent

Prior to inclusion, patients gave their informed written consent to participate in the study. The study was conducted according to International Standards of Good Clinical Practice (ICH guidelines and the Declaration of Helsinki). The project was approved by the local Research Ethics Committees of San Juan de Dios Hospital, Santiago, Chile.

RESULTS

Demographics, Clinical Variables, and A-WBAR (Table 1)

There were no significant differences in age [IPD: 62.2 ± 11.5 (years \pm SD); PSP 68.2 ± 6.3 ; MSA 60.4 ± 7.7], gender, and disease duration between the IPD patients and both the MSA and PSP patients, although PSP patients were significantly older than MSA patients ($p < 0.001$). Disease duration was <5 years for all groups. The MSA patients had a longer disease duration with a mean of 4.3 years [PSP 3.0 ($p = 0.04$)]. The PSP and MSA patients showed greater impairment on the H&Y scale than the IPD patients. The PSP patients showed greater impairment on the cognitive measures than the IPD and MSA patients.

MSA and PSP, but not IPD, showed significant mean deterioration over the follow-up period on a range of clinical measures.

In controls, the a-WBAR was $0.37\% \pm 0.28$ (CI 95% 0.17–0.57), while in IPD patients a-WBAR was $0.54\% \pm 0.38$ (CI 95% 0.32–0.68). The IPD patients did not differ from the controls. In PSP patients, the a-WBAR was $1.93\% \pm 1.1$ (CI 95% 1.5–2.2). In MSA patients, a-WBAR was $1.65\% \pm 0.9$ (CI 95% 1.37–1.93). The MSA group did not differ from the PSP group. a-WBAR in the PSP and MSA groups was significantly higher than in the IPD group ($p < 0.001$; **Figure 1**). a-WBAR 0.6% differentiated patients with IPD from those with PSA and MSA with 91% sensitivity and 80% specificity (**Figure 2**); for IPD vs. MSA groups this value shows 85% sensitivity and 80% specificity, and for IPD vs. PSP groups 97% sensitivity and 75% specificity.

DISCUSSION

Diagnosis of Parkinsonian syndromes remains a difficult task that is based mainly on the clinical evaluation of neurologists, as no biological markers are currently available (Adler et al., 2014). Misdiagnosis not only means that patients may suffer from prognostic uncertainty but also means that clinical investigations

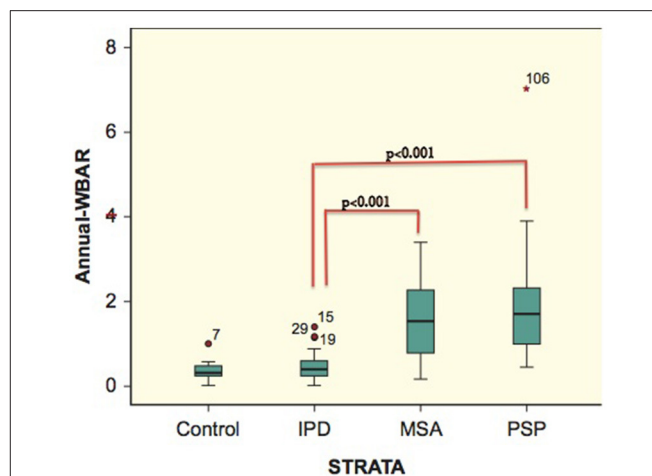


FIGURE 1 | Error bars showing 95% confident intervals (bars) of a-WBAR means for each group. 7, 15, 29, 19, and 106 = outliers. STRATA = group.

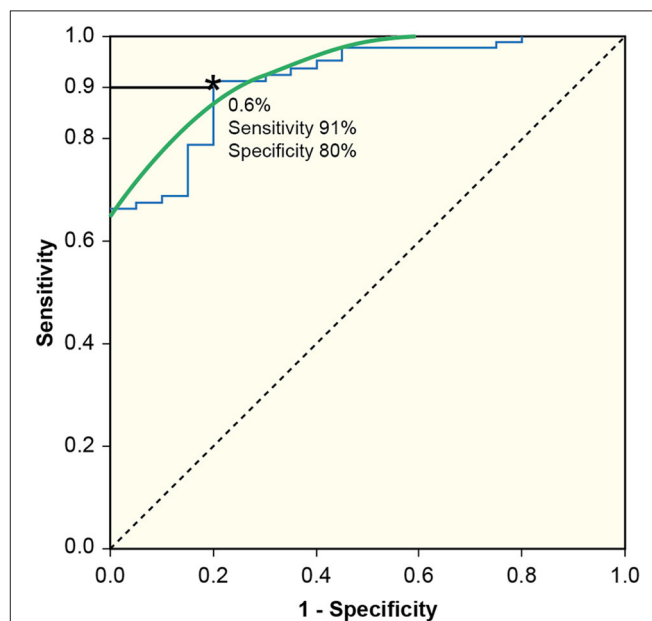


FIGURE 2 | Receiver Operating Characteristic curve (ROC) to define maximal sum of sensitivity and specificity. a-WBAR at 0.6% is the cut-off point and differentiated patients with IPD from those with PSP and MSA with 91% sensitivity and 80% specificity.

are hampered by false positive cases. The inaccuracy in diagnosis is explained by the unknown *tempo* of widespread cellular destruction and the variable sites within the nigro-striatal dopaminergic system and/or cortices where neurodegeneration commences. A recent clinicopathologic study indicated that the clinical diagnostic capabilities for IPD have not advanced over the last 23 years (Rajput et al., 1991; Hughes et al., 1993; Adler et al., 2014), with only 26% accuracy for the clinical diagnosis of untreated or not clearly responsive patients, and 53% accuracy in

patients who respond early to medications with disease duration <5 years (Adler et al., 2014). An early longitudinal diagnostic biomarker to help to differentiate IPD vs. MSA and PSP is still a needed and was the main aim of this study.

In this study IPD patients did not show abnormal a-WBAR as was previously reported for IPD without dementia [0.6% (Paviour et al., 2006), 0.28% (Burton et al., 2005)]. We found an a-WBAR of 1.93% for PSP and 1.65% for MSA, which are higher than those in previous reports: approximately 1% for both PSP and MSA using BSI (Josephs et al., 2006; Paviour et al., 2006; Whitwell et al., 2007). Consistent with those reports, in the current study no significant difference was observed between a-WBAR in PSP and MSA (Paviour et al., 2006).

The a-WBAR reported in MSA and PSP are somewhat closer to those reported for Alzheimer disease using both BSI [2.1% (Schott et al., 2005), 2.37% (Chan et al., 2001), 2.78% (Fox and Freeborough, 1997)] and SIENA [1.9% (Sluiter et al., 2008b)]. It is plausible that cortical structures are the main contributors to whole brain atrophy in PSP and MSA. In PSP, neuronal loss is recognized in frontal, temporal, and limbic cortices and much less in parietal and occipital cortices (Verny et al., 1996). Such a neuronal loss is not considered to be typical in MSA. However, Papp and Lantos described high densities of glial cytoplasmic inclusions in the supplementary and primary motor cortical areas and subadjacent white matter and moderate densities of glial cytoplasmic inclusions in the premotor area, cingulate motor area, and corpus callosum in MSA (Papp and Lantos, 1994). In a review of 203 proven MSA cases, some degree of cortical atrophy was observed in 21% of cases (Wenning et al., 1997), and *post mortem* examinations showed severe frontal atrophy (Inoue et al., 1997; Wakabayashi et al., 1998). *In vivo* data in MSA showed hypometabolism in motor, premotor and prefrontal cortices and parietal lobes (Kawai et al., 2008). A proton magnetic resonance spectroscopy study showed a significant reduction of N-acetylaspartate/creatine in the frontal cortex (Abe et al., 2000). Voxel-based morphometry studies have suggested that atrophy in the motor and prefrontal cortices are common findings in MSA (Brenneis et al., 2003).

By contrast, in levodopa-responsive IPD patients, evidence supports the idea that motor deficits are primarily related to the localized loss of selective dopaminergic neurons in the substantia nigra, with cortical and subcortical gray and white matter structures more preserved in comparison with those with PSP and MSA.

From a clinical perspective, an a-WBAR cutoff point of 0.6% may provide a potential retrospective application for a-WBAR to improve diagnostic accuracy (91% sensitivity and 80% specificity) for IPD vs. PSP and MSA, particularly in the initial stages when the clinical “plus syndrome” has not yet manifested and the response to levodopa treatment is being assessed.

With the current limited knowledge about the biology of MSA and PSP, interpretations and designs of MRI studies are mainly based on the information provided by proven cases (region of interest based studies). Sensitivity and specificity have been reported for many neuroimaging techniques based on region of interests for the differential diagnosis of IPD vs. MSA and PSP. Metabolic imaging using positron emission

tomography (PET) studies of glucose metabolism were reported to have 86% sensitivity and 91% specificity to correctly categorize IPD from MSA and PSP (Hellwig et al., 2012). Dopamine transport (DAT) imaging using single photon emission CT (DAT-SPECT) is not efficient for the differentiation of IPD from PSP and MSA (Lokkegaard et al., 2002). Both molecular techniques PET and DAT-SPECT are expensive and not routinely available. A diffusion weighted imaging (DWI) study reported 90% sensitivity for differentiating PSP from IPD; however, in this study DWI was evaluated in only 10 PSP and 13 IPD patients (Seppi et al., 2003). Transcranial sonography has been reported to have 40% sensitivity and 61% specificity for the diagnosis of IPD (Bouwman et al., 2013). Considering these data, a review concludes that no techniques are specifically recommended for routine use in clinical practice (Politis, 2014).

For disease-modifying treatments, the current challenge is to find biomarkers to accurately differentiate IPD from the aggressive MSA and PSP, early in the disease course. Ideally, MRI studies should also be based, as much as possible, on information obtained during the natural course of these diseases. The clinical and pathological aggressiveness of MSA and PSP may be due to global brain atrophy rather than degeneration of specific brain pathways and/or gray matter structures. a-WBAR within a normal range is unlikely to be observed in PSP or MSA but is likely to be observed in IPD patients. We propose a complementary use of clinical features (bradykinesia, rigidity, resting tremor, and response to dopaminergic drugs) and a-WBAR as a reasonable approach for the most accurate clinical diagnosis in these disorders early in the disease course.

A problem with using brain volume as a disease outcome is that it may not reflect physiologic or synaptic health. Furthermore, loss of brain volume might be influenced by causes that are common in people with chronic brain disorders, but only indirectly related to the disease itself, such as mild traumatic brain injury (MacKenzie et al., 2002), chronic alcohol abuse (Bartsch et al., 2007), nutritional deficiency, or hydration/dehydration (Kempton et al., 2009). However, these sources of variation are certainly less than that due to the disease itself.

As the current state of the art technique in neuroimaging, SIENA may be among the simplest MRI tools, but complex methodologies do not necessarily lead to robust and coherent results (Smith et al., 2004).

Overall, this study supports a complementary use of clinical tools and global rates of brain atrophy as an aid to clinical diagnosis between IPD vs. PSP and MSA.

AUTHOR CONTRIBUTIONS

CG: Research project: Conception, Organization and Execution. Statistical Analysis: Design, Execution, Review and Critique. Manuscript: Writing of the first draft, Review and Critique. KB: Research project: Organization and Execution. Manuscript: Review and Critique. WS: Research

project: Organization and Execution. Manuscript: Review and Critique. GG: Research project: Organization and Execution. Manuscript: Review and Critique. GF: Research project: Organization and Execution. Manuscript: Review and Critique.

REFERENCES

- Abe, K., Terakawa, H., Takanashi, M., Watanabe, Y., Tanaka, H., Fujita, N., et al. (2000). Proton magnetic resonance spectroscopy of patients with parkinsonism. *Brain Res. Bull.* 52, 589–595. doi: 10.1016/S0361-9230(00)00321-X
- Adler, C. H., Beach, T. G., Hentz, J. G., Shill, H. A., Caviness, J. N., Driver-Dunckley, E., et al. (2014). Low clinical diagnostic accuracy of early vs. advanced Parkinson disease: clinicopathologic study. *Neurology* 83, 406–412. doi: 10.1212/WNL.0000000000000641
- Bartsch, A. J., Homola, G., Biller, A., Smith, S. M., Weijers, H. G., Wiesbeck, G. A., et al. (2007). Manifestations of early brain recovery associated with abstinence from alcoholism. *Brain* 130, 36–47. doi: 10.1093/brain/awl303
- Bouwman, A. E. P., Vlaar, A. M. M., Mess, W. H., Kessels, A., and Weber, W. E. J. (2013). Specificity and sensitivity of transcranial sonography of the substantia nigra in the diagnosis of Parkinson's disease: prospective cohort study in 196 patients. *BMJ Open* 3:e002613. doi: 10.1136/bmjopen-2013-002613
- Brenneis, C., Seppi, K., Schocke, M. F., Müller, J., Luginger, E., Bösch, S., et al. (2003). Voxel-based morphometry detects cortical atrophy in the Parkinson variant of multiple system atrophy. *Mov. Disord.* 18, 1132–1138. doi: 10.1002/mds.10502
- Burton, E. J., McKeith, I. G., Burn, D. J., and O'Brien, J. T. (2005). Brain atrophy rates in Parkinson's disease with and without dementia using serial magnetic resonance imaging. *Mov. Disord.* 20, 1571–1576. doi: 10.1002/mds.20652
- Chan, D., Fox, N. C., Jenkins, R., Scallan, R. I., Crum, W. R., and Rossor, M. N. (2001). Rates of global and regional cerebral atrophy in AD and frontotemporal dementia. *Neurology* 57, 1756–1763. doi: 10.1212/WNL.57.10.1756
- Dubois, B., Slachevsky, A., Litvan, I., and Pillon, B. (2000). The FAB: A Frontal Assessment Battery at bedside. *Neurology* 55, 1621–1626. doi: 10.1212/WNL.55.11.1621
- Fox, N. C., and Freeborough, P. A. (1997). Brain atrophy progression measured from registered serial MRI: validation and application to Alzheimer's disease. *J. Magn. Reson. Imaging* 7, 1069–1075. doi: 10.1002/jmri.1880070620
- Freeborough, P. A., Fox, N. C., and Kitney, R. I. (1997). Interactive algorithms for the segmentation and quantitation of 3-D MRI brain scans. *Comput. Methods Programs Biomed.* 53, 15–25. doi: 10.1016/S0169-2607(97)01803-8
- Goetz, C. G., Fahn, S., Martinez-Martin, P., Poewe, W., Sampaio, C., Stebbins, G. T., et al. (2007). Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): Process, format, and clinimetric testing plan. *Mov. Disord.* 22, 41–47. doi: 10.1002/mds.21198
- Gordon, E., Rohrer, J. D., Kim, L. G., Omar, R., Rossor, M. N., Fox, N. C., et al. (2010). Measuring disease progression in frontotemporal lobar degeneration: a clinical and MRI study. *Neurology* 74, 666–673. doi: 10.1212/WNL.0b013e3181d1a879
- Hellwig, S., Amtege, F., Kreft, A., Buchert, R., Winz, O. H., Vach, W., et al. (2012). [(1)(8)F]FDG-PET is superior to [(1)(2)(3)I]IBZM-SPECT for the differential diagnosis of parkinsonism. *Neurology* 79, 1314–1322. doi: 10.1212/WNL.0b013e31826c1b0a
- Hobbs, N. Z., Barnes, J., Frost, C., Henley, S. M., Wild, E. J., Macdonald, K., et al. (2010). Onset and progression of pathologic atrophy in Huntington disease: a longitudinal MR imaging study. *AJNR Am. J. Neuroradiol.* 31, 1036–1041. doi: 10.3174/ajnr.A2018
- Hoehn, M. M., and Yahr, M. D. (1967). Parkinsonism: onset, progression, and mortality. *Neurology* 17, 427–442.
- Hughes, A. J., Daniel, S. E., Blankson, S., and Lees, A. J. (1993). A clinicopathologic study of 100 cases of Parkinson's disease. *Arch. Neurol.* 50, 140–148. doi: 10.1001/archneur.1993.00540020018011
- Hughes, A. J., Daniel, S. E., and Lees, A. J. (2001). Improved accuracy of clinical diagnosis of Lewy body Parkinson's disease. *Neurology* 57, 1497–1499. doi: 10.1212/WNL.57.8.1497
- Inoue, M., Yagishita, S., Ryo, M., Hasegawa, K., Amano, N., and Matsushita, M. (1997). The distribution and dynamic density of oligodendroglial cytoplasmic inclusions (GCIs) in multiple system atrophy: a correlation between the density of GCIs and the degree of involvement of striatonigral and olivopontocerebellar systems. *Acta Neuropathol.* 93, 585–591. doi: 10.1007/s004010050655
- Josephs, K. A., Whitwell, J. L., Boeve, B. F., Shiung, M. M., Gunter, J. L., Parisi, J. E., et al. (2006). Rates of cerebral atrophy in autopsy-confirmed progressive supranuclear palsy. *Ann. Neurol.* 59, 200–203. doi: 10.1002/ana.20707
- Kawai, Y., Suenaga, M., Takeda, A., Ito, M., Watanabe, H., Tanaka, F., et al. (2008). Cognitive impairments in multiple system atrophy: MSA-C vs. MSA-P. *Neurology* 70, 1390–1396. doi: 10.1212/01.wnl.0000310413.04462.6a
- Kempton, M. J., Ettinger, U., Schmechtig, A., Winter, E. M., Smith, L., McMorris, T., et al. (2009). Effects of acute dehydration on brain morphology in healthy humans. *Hum. Brain Mapp.* 30, 291–298. doi: 10.1002/hbm.20500
- Litvan, I., Bhatia, K. P., Burn, D. J., Goetz, C. G., Lang, A. E., McKeith, I., et al. (2003). Movement Disorders Society Scientific Issues Committee report: SIC Task Force appraisal of clinical diagnostic criteria for Parkinsonian disorders. *Mov. Disord.* 18, 467–486. doi: 10.1002/mds.10459
- Lokkegaard, A., Werdelin, L. M., and Friberg, L. (2002). Clinical impact of diagnostic SPET investigations with a dopamine re-uptake ligand. *Eur. J. Nucl. Med. Mol. Imaging* 29, 1623–1629. doi: 10.1007/s00259-002-0938-7
- Mackenzie, J. D., Siddiqi, F., Babb, J. S., Bagley, L. J., Mannon, L. J., Sinson, G. P., et al. (2002). Brain atrophy in mild or moderate traumatic brain injury: a longitudinal quantitative analysis. *AJNR Am. J. Neuroradiol.* 23, 1509–1515.
- Papp, M. I., and Lantos, P. L. (1994). The distribution of oligodendroglial inclusions in multiple system atrophy and its relevance to clinical symptomatology. *Brain* 117 (Pt 2), 235–243. doi: 10.1093/brain/117.2.235
- Paviour, D. C., Price, S. L., Jahanshahi, M., Lees, A. J., and Fox, N. C. (2006). Longitudinal MRI in progressive supranuclear palsy and multiple system atrophy: rates and regions of atrophy. *Brain* 129, 1040–1049. doi: 10.1093/brain/awl021
- Politis, M. (2014). Neuroimaging in Parkinson disease: from research setting to clinical practice. *Nat. Rev. Neurol.* 10, 708–722. doi: 10.1038/nrneurol.2014.205
- Rajput, A. H., Rozdilsky, B., and Rajput, A. (1991). Accuracy of clinical diagnosis in parkinsonism—a prospective study. *Can. J. Neurol. Sci.* 18, 275–278. doi: 10.1017/S0317167100031814
- Respondek, G., Stamelou, M., Kurz, C., Ferguson, L. W., Rajput, A., Chiu, W. Z., et al. (2014). The phenotypic spectrum of progressive supranuclear palsy: a retrospective multicenter study of 100 definite cases. *Mov. Disord.* 29, 1758–1766. doi: 10.1002/mds.26054
- Ridha, B. H., Anderson, V. M., Barnes, J., Boyes, R. G., Price, S. L., Rossor, M. N., et al. (2008). Volumetric MRI and cognitive measures in Alzheimer disease: comparison of markers of progression. *J. Neurol.* 255, 567–574. doi: 10.1007/s00415-008-0750-9
- Schott, J. M., Price, S. L., Frost, C., Whitwell, J. L., Rossor, M. N., and Fox, N. C. (2005). Measuring atrophy in Alzheimer disease: a serial MRI study over 6 and 12 months. *Neurology* 65, 119–124. doi: 10.1212/01.wnl.0000167542.89697.0f
- Seppi, K., Schocke, M. F., Esterhammer, R., Kremser, C., Brenneis, C., Mueller, J., et al. (2003). Diffusion-weighted imaging discriminates progressive supranuclear palsy from PD, but not from the parkinson variant of multiple system atrophy. *Neurology* 60, 922–927. doi: 10.1212/01.WNL.0000049911.91657.9D
- Suimer, J. D., Van Der Flier, W. M., Karas, G. B., Fox, N. C., Scheltens, P., Barkhof, F., et al. (2008a). Whole-brain atrophy rate and cognitive decline: longitudinal MR study of memory clinic patients. *Radiology* 248, 590–598. doi: 10.1148/radiol.2482070938
- Suimer, J. D., Vrenken, H., Blankenstein, M. A., Fox, N. C., Scheltens, P., Barkhof, F., et al. (2008b). Whole-brain atrophy rate in Alzheimer disease: identifying fast progressors. *Neurology* 70, 1836–1841. doi: 10.1212/01.wnl.0000311446.61861.e3

FUNDING

This study was supported by FONDECYT, grant 11121212, from the Chilean government and OAIC Hospital Clinico de la Universidad de Chile.

- Smith, S. M., De Stefano, N., Jenkinson, M., and Matthews, P. M. (2001). Normalized accurate measurement of longitudinal brain change. *J. Comput. Assist. Tomogr.* 25, 466–475. doi: 10.1097/00004728-200105000-00022
- Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E. J., Johansen-Berg, H., et al. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage* 23(Suppl. 1), S208–S219. doi: 10.1016/j.neuroimage.2004.07.051
- Smith, S. M., Zhang, Y., Jenkinson, M., Chen, J., Matthews, P. M., and Federico, A., et al (2002). Accurate, robust, and automated longitudinal and cross-sectional brain change analysis. *Neuroimage* 17, 479–489. doi: 10.1006/nimg.2002.1040
- Verny, M., Duyckaerts, C., Agid, Y., and Hauw, J. J. (1996). The significance of cortical pathology in progressive supranuclear palsy. Clinico-pathological data in 10 cases. *Brain* 119 (Pt 4), 1123–1136. doi: 10.1093/brain/119.4.1123
- Wakabayashi, K., Ikeuchi, T., Ishikawa, A., and Takahashi, H. (1998). Multiple system atrophy with severe involvement of the motor cortical areas and cerebral white matter. *J. Neurol. Sci.* 156, 114–117. doi: 10.1016/S0022-510X(98)00018-5
- Wenning, G. K., Tison, F., Ben Shlomo, Y., Daniel, S. E., and Quinn, N. P. (1997). Multiple system atrophy: a review of 203 pathologically proven cases. *Mov. Disord.* 12, 133–147. doi: 10.1002/mds.870120203
- Whitwell, J. L., Jack, C. R. Jr., Parisi, J. E., Knopman, D. S., Boeve, B. F., Petersen, R. C., et al. (2007). Rates of cerebral atrophy differ in different degenerative pathologies. *Brain* 130, 1148–1158. doi: 10.1093/brain/awm021

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.

Copyright © 2017 Guevara, Bulatova, Soruco, Gonzalez and Farias. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Advantages of publishing in Frontiers



OPEN ACCESS

Articles are free to read
for greatest visibility
and readership



FAST PUBLICATION

Around 90 days
from submission
to decision



HIGH QUALITY PEER-REVIEW

Rigorous, collaborative,
and constructive
peer-review



TRANSPARENT PEER-REVIEW

Editors and reviewers
acknowledged by name
on published articles

Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org

Contact us: info@frontiersin.org | +41 21 510 17 00



REPRODUCIBILITY OF RESEARCH

Support open data
and methods to enhance
research reproducibility



DIGITAL PUBLISHING

Articles designed
for optimal readership
across devices



FOLLOW US

@frontiersin



IMPACT METRICS

Advanced article metrics
track visibility across
digital media



EXTENSIVE PROMOTION

Marketing
and promotion
of impactful research



LOOP RESEARCH NETWORK

Our network
increases your
article's readership