

Université de Montréal

**Selection of multiscale functional brain connections for early diagnosis of
Alzheimer's disease in multicentric studies**

par
Christian Dansereau

Département d'informatique et de recherche opérationnelle
Faculté des arts et des sciences

Rapport pour la partie orale
de l'examen pré-doctoral

Décembre, 2014

© Christian Dansereau, 2014.

Université de Montréal
Faculté des études supérieures

Cet examen pré-doctoral intitulé:

**Selection of multiscale functional brain connections for early diagnosis of
Alzheimer's disease in multicentric studies**

présenté par:

Christian Dansereau

a été évalué par un jury composé des personnes suivantes:

Jean Meunier, président-rapporteur
Pierre Bellec, directeur de recherche
Aaron Courville, membre du jury

Examen accepté le:

ABSTRACT

Functional magnetic resonance imaging (fMRI) gives a non-invasive measure of functional connectivity throughout the brain in individual human subjects. This imaging modality captures a massive amount of connections, in the order of 10^8 , some of which may be biomarkers of Alzheimer's disease (AD). The brain is highly structured into a nested hierarchy of networks, which can be leveraged to reduce the dimensionality of brain connectivity into a limited set of biologically meaningful features. The objective of this project is to develop multiscale clustering techniques for feature extraction, in conjunction with random forest for prediction of clinical diagnosis and prognosis of AD, using fMRI connectivity. I will investigate the case of multicentric fMRI data including a few hundreds of participants, i.e. the AD neuroimaging initiative (ADNI) sample, which is the typical design found in phase II pharmaceutical clinical trials. Site-specific MRI set-ups may bias the fMRI measures, and I am thus developing procedures for inter-site normalization. The main outcome of this project will be a prediction pipeline for the data-driven identification of AD biomarkers in resting-state fMRI.

Keywords: **fMRI, Alzheimer's disease, biomarker, multisite.**

CONTENTS

ABSTRACT	iii
CONTENTS	iv
LIST OF TABLES	vi
LIST OF FIGURES	vii
CHAPTER 1: GENERAL CONTEXT	1
1.1 Alzheimer's disease	2
1.2 Overview of functional magnetic resonance imaging	5
1.3 Preprocessing	6
1.4 Multi-site	8
1.5 Resting-state connectivity	11
1.6 Prediction	16
1.6.1 Prediction in the context of AD	16
1.6.2 Importance of preprocessing to improve classification	17
1.6.3 High dimensionality problem	17
1.6.4 Feature selection	18
1.6.5 Classification	18
1.7 Objectives	19
CHAPTER 2: SCRUBBING AND MOTION ARTEFACT CORRECTION	20
2.1 Introduction	20
2.2 Method	22
2.2.1 Dataset	22
2.2.2 Preprocessing	22
2.2.3 Statistical analysis	23
2.3 Results	24

2.3.1	Motion distribution in different clinical cohorts	24
2.3.2	Default mode network in young adult and elderly population . .	26
2.3.3	Impact of scrubbing on the connectivity	28
2.3.4	Impact of scrubbing on our discriminative power between populations	34
2.4	Conclusion	36
CHAPTER 3: FEASIBILITY OF MULTI-CENTRIC FMRI CONNECTIVITY STUDIES OF ALZHEIMER’S DISEASE		37
3.1	Introduction	37
3.2	Method	37
3.2.1	Preprocessing	37
3.2.2	Feature selection	38
3.2.3	Simulations	42
3.3	Results	44
3.4	Conclusion	54
CHAPTER 4: FUTURE DEVELOPMENT		55
4.1	The connectome feature space	55
4.2	Feature selection and classification	56
4.3	Industry application and translational effort	58
4.4	Timeline	58
REFERENCES		61

LIST OF TABLES

2.I	Retention table after scrubbing	25
3.I	Regions selected in the literature review, the region number correspond to the number in the Cambridge 100 partition.	49

LIST OF FIGURES

1.1	Biomarker model of Alzheimer's disease	4
1.2	Schematic of the BOLD effect	6
1.3	Schematic of the preprocessing	7
1.4	Motion estimation	8
1.5	Resting-state correlation maps	12
1.6	Functional connectome	13
1.7	Consistent resting-state network (CRSN)	15
2.1	FD distribution among groups	26
2.2	Preprocessing impact on DMN for CNY and CNE	27
2.3	Impact of the preprocessing on the DMN (CNE,pMCI,pDAT) . .	29
2.4	Impact of the preprocessing on the DMN (CNY)	30
2.5	Impact of preprocessing on each group	31
2.6	Scrubbing impact on group differences	32
2.7	Scrubbing impact on group differences	33
2.8	Scrubbing impact on group differences	33
2.9	Detection power with Scrubbing	35
2.10	Detection power with CompCor and GSC	36
3.1	ICC scores	45
3.2	Cited region frequency	46
3.3	Selected region inside DMN	47
3.4	Selected region outside DMN	48
3.5	inter vs. intra site variability	50
3.6	DMN variability across sites	51
3.7	Detection power	52
3.8	Detection power 2	53

CHAPTER 1

GENERAL CONTEXT

The number of Canadians suffering from Alzheimer's disease (AD) is rapidly increasing, with tremendous social and economic impact. Despite the emergence of promising drugs, the recent clinical trials with demented patients have failed. Dementia however comes very late in the development of the disease, at a stage where the degeneration of neural tissues has likely gone beyond repair. In order to be efficient, therapies should be initiated in the decades predating dementia, in a preclinical stage where patients experience no or very mild symptoms (see chapter 1.1). There are unfortunately no biomarker(s) that are currently predictive of AD in this preclinical stage, and could help identify the individuals that could benefit from such interventions. A promising technique is resting-state functional magnetic resonance imaging (rs-fMRI), which may be able to capture the early synaptic dysfunction seen in AD (see chapter 1.2). In order to be able to apply statistical analysis and machine learning methods we need to preprocess the data to remove as much as possible the effect of various artefacts (hardware and physiological). The preprocessing reduces the variability of the data and therefore provides more relevant and discriminative features (see chapter 1.3). In order to further improve the statistical power of there analysis, multiple academic groups are collaborating to pool their dataset to increase the sample size of the study. Unfortunately the gain in sample size comes with a new source of variability introduced by the multicentric acquisition. Site-specific MRI set-ups may bias the fMRI measures, and I am thus developing procedures for inter-site normalization. Account for these sources of variance are important since they may bias the predictive potential of rs-functional connectivity in a multi-site, in line with this last assumption we want to quantify the robustness of the feature selection and classifier to multi-site acquisition (see chapter 1.4).

With the objective of extracting meaningful information that could be used as biomarkers we are using connectivity measures usually based on the correlation of spontaneous fluctuations in neurovascular activity from pairs of brain regions, which have been shown

to be sensitive to the development of the disease (see chapter 1.5). However, with about 10^4 recording sites in the gray matter and close to no knowledge on the early brain dysfunction in AD, there is an overwhelming number of 10^7 possible connections to examine as a potential diagnostic we will therefore focus on the feature selection procedure and the stability of those selections in order to have consistent and sparse predictive features discriminative of the disease(see chapter 1.6). The main outcome of this project will be a prediction pipeline for the data-driven identification of AD biomarkers in resting-state fMRI.

1.1 Alzheimer's disease

Alzheimer's disease (AD) is a major neurodegenerative disorder characterized by cognitive and intellectual deficits and behavioural changes without a known cause or an effective treatment. It gradually destroys a patient's memory and ability to reason, make judgments, communicate and carry out daily activities (Jeong 2004). With the aging of the population worldwide, this disorder has attracted much attention. Evidence from clinical elderly individuals suggests that the pathophysiological process of AD begins years, if not decades, before the diagnosis of clinical dementia (Morris 2005). The clinical disease stages of AD are divided into three phases described by Jack and colleagues Jack et al. (2010).

First is a pre-symptomatic phase in which individuals are cognitively normal but some have pathological changes in AD. Second is a prodromal phase of AD, commonly referred to as mild cognitive impairment (MCI) (Petersen 2004), which is characterised by the onset of the earliest cognitive symptoms (typically deficits in episodic memory) that do not meet the criteria for dementia. The severity of cognitive impairment in the MCI phase of AD varies from early manifestation of memory dysfunction to more widespread dysfunction in other cognitive domains. The final phase in the evolution of AD is dementia, defined as multi-domain impairments that are severe enough to result in loss of function (Jack et al. 2010).

The use of biomarker for the early diagnostic of pathologies has a long history, with

many studies showing the feasibility of using AD biomarker to predict conversion from MCI to AD. These studies show that individuals on the course of developing AD can be identified earlier in the course of the disease by using the MCI stage with the addition of imaging and cerebrospinal fluid (CSF) biomarkers to enhance diagnostic specificity (Chetelat et al. 2003, Jack et al. 1999, Mattsson et al. 2009, Yuan et al. 2009). It could be possible to diagnose AD after the exclusion of other forms of dementia, although a formal diagnostic can only be made after a post-mortem evaluation of the brain tissue (McKhann et al. 1984). This is one of the reasons why MRI based analysis and diagnostic tools are currently a hot topic in clinical neuroscience research.

The currently dominant hypothesis in the field for the chain of events in AD pathophysiology is the β -amyloid ($A\beta$)-cascade hypothesis. It suggests that interstitial $A\beta$ proteins exert a toxic effect on surrounding neurons and synapses by forming plaques, thereby disturbing their function (Hardy and Selkoe 2002, Shankar et al. 2008). Moreover, recent research study suggests that, prior to neuronal death resulting in brain atrophy, disruption of functional connectivity may arise in response to an unknown systemic problem and represent an early outcome of $A\beta$ proteins plaque formation in AD (Sheline and Raichle 2013). Atrophy is the result of neuronal death and is measures in vivo using structural MRI measuring the thickness of the gray matter on the cortex (also called cortical thickness). Already in the stage preceding aggregation of $A\beta$ fragments into amyloid plaques, there is a dysfunction of synaptic transmission in many brain regions due to dimers and monomers from the $A\beta$ cascade (D'Amelio and Rossini 2012). As illustrated in Sperling et al. (2011) a viable hypothesis is that fMRI precedes the structural changes as well as clinical symptoms and is believed to start in the preclinical phase of the disease.

R.A. Sperling et al. / Alzheimer's & Dementia 7 (2011) 280–292

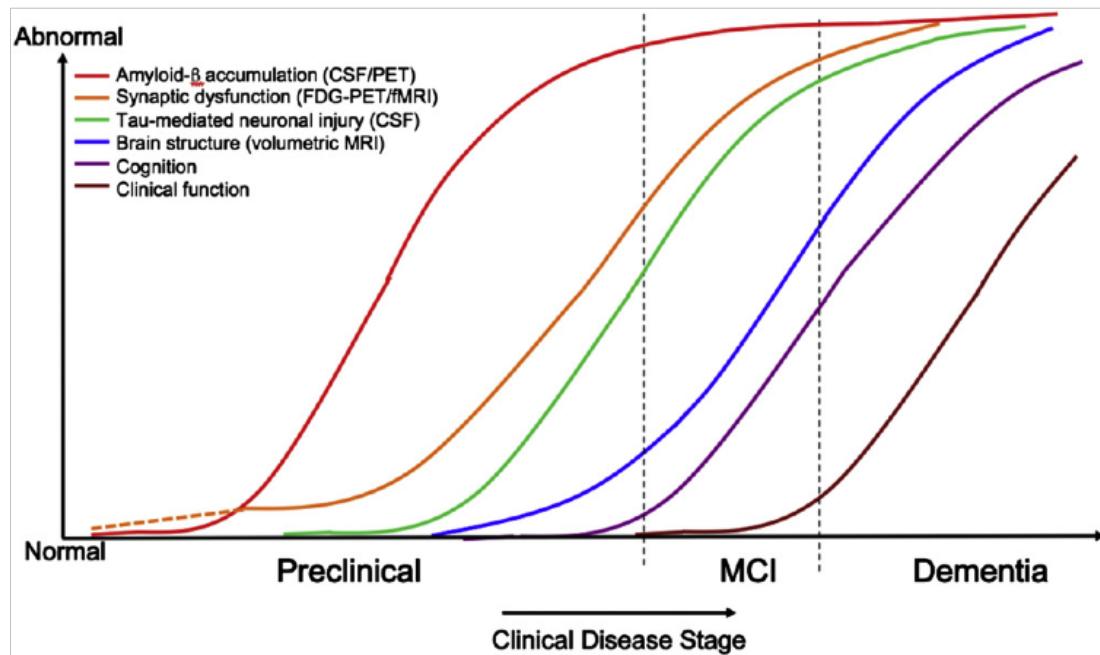


Figure 1.1: Hypothetical model of dynamic biomarkers of the AD expanded to explicate the preclinical phase: A β as identified by cerebrospinal fluid A β 42 assay or PET amyloid imaging. Synaptic dysfunction evidenced by fluorodeoxyglucose (F18) positron emission tomography (FDG-PET) or functional magnetic resonance imaging (fMRI), with a dashed line to indicate that synaptic dysfunction may be detectable in carriers of the 34 allele of the apolipoprotein E gene before detectable A β deposition. Neuronal injury is evidenced by cerebrospinal fluid tau or phospho-tau, brain structure is evidenced by structural magnetic resonance imaging. Biomarkers change from normal to maximally abnormal (y-axis) as a function of disease stage (x-axis). The temporal trajectory of two key indicators used to stage the disease clinically, cognitive and behavioral measures, and clinical function are also illustrated. Figure from Sperling et al. (2011).

1.2 Overview of functional magnetic resonance imaging

In functional magnetic resonance imaging (fMRI), the acquisition process is slightly different than for the anatomical MRI acquisition. fMRI uses the principle of the relaxation of hydrogen nuclei, by using specific fMRI sequences consisting in $T2^*$ weighted acquisitions that are sensitive to local distortions of the magnetic field. Particularly, deoxyhemoglobin (a form of hemoglobin without oxygen) will create such local distortions of the magnetic field, since it is a paramagnetic molecule (positive magnetic susceptibility) (Ogawa et al. 1990). The data acquired using fMRI rely on the hypothesis that areas showing decreased deoxyhemoglobin concentration are due to sustained brain activity. Following neuronal activity, neurons require energy to restore the electrical and ionic concentration balance across the cell membrane. One mechanism to generate this energy is the glucose oxidative metabolism, which requires the delivery of oxygen and glucose by the blood on the site where brain activity takes place (Ogawa et al. 1990) . Initially, fMRI was thought to be a good technique to measure the cerebral metabolic rate of oxygen, since the new blood rushing in causes a proportional effect on the venous-end, resulting in a decrease in deoxyhemoglobin concentration, and thus an increase in fMRI signal. The concentration of deoxyhemoglobin actually depends mainly on three factors or phenomena: metabolic rate of oxygen consumption, cerebral blood volume (CBV) and cerebral blood flow (CBF) (Hoge et al. 1999). As a result, the fMRI signal is the outcome of competing effects following neuronal activity and fortunately these effects do not cancel out, allowing the detection of a net signal increase. We call the interrelation of these effects the blood oxygenation level dependent (BOLD) effect.

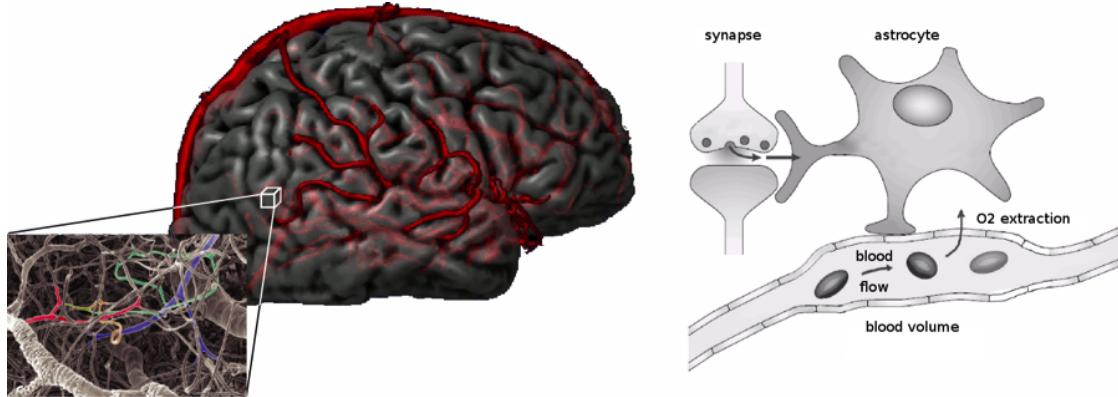


Figure 1.2: Representation of the brain and its vasculature (on the left) and a schematic view of the interaction between the effect of neuronal activity on local changes in blood oxygenation signal (BOLD) (on the right) (adapted from Heeger and Ress (2002)).

1.3 Preprocessing

Normalization of the data is crucial to obtain a consistent and accurate classifier (Kotsiantis 2007). Therefore a particular attention is placed on the correction and normalization procedure applied to the rs-fMRI data used in this study. A series of standard preprocessing steps is usually applied in an attempt to correct for various artefacts that would perturb the subsequent analysis. The BOLD effect associated with neuronal activity generally results in a relatively small fluctuation of the MR signal. Many factors can influence this signal. Among them, the physiological activity associated mainly with respiration, cardiac pulsations and patient's motion are major contributors to the noise and are spatially spread everywhere within the brain volume. These sources of noise result in large correlations between BOLD signals of distant voxels. An other factor is the fact that we need a form of spatial normalization of the individual brains in order to perform analysis across subjects (due to anatomical variance among subjects), this spatial normalization (coregistration of the individual brains with a reference template) is necessary but can potentially be an other source of confound. Therefore, preprocessing methods were designed in an attempt to remove specifically the so-called structured

noise and motion artefacts from the raw fMRI data. A schematic representation of the preprocessing pipeline can be seen in Figure 1.3.

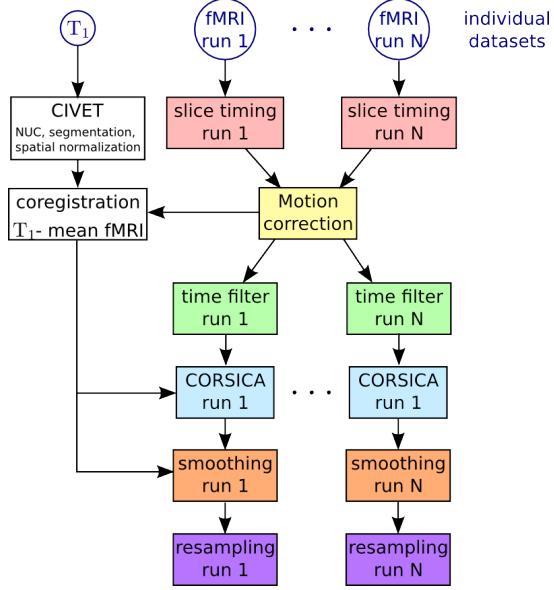


Figure 1.3: Schematic of the preprocessing pipeline including spatial and functional normalization (NIAK preprocessing pipeline ²).

The basic steps are as follow: (1) correction for slice timing differences due to delay in acquisition sampling; (2) rigid-body motion estimation for within and between runs, motion correction operates by selecting one functional volume as a reference to align all other functional volumes. Most head motion algorithms describe head movement by 6 parameters, three translation (displacement) parameters and rotation parameters and are appropriate to characterize motion of rigid bodies (see Figure 1.4); (3) Coregistration of the functional data in a reference space; (4) resampling of the functional data in the stereotaxic space (references brain used as a common space between subjects); (5) regression of confounds in order to remove spatially structured noise on the fMRI time-series. The confounds are the slow time drift, high frequency noise signal, motion parameters, the average signal white matter as well as the average signal of the ventricles (containing cerebrospinal fluid CSF a frequent source of noise and artefact). Some

groups have suggested that these corrections are not sufficient to remove motion artefact and propose some additional corrective procedure (detailed in Chapter 2); and (6) the spatial smoothing is usually applied using a Gaussian blurring kernel to improve signal to noise ratio (SNR), improved validity of the statistical tests by making the error distribution more normal and finally reduce anatomical and functional variations between subjects (Mikl et al. 2008, Worsley and Friston 1995).

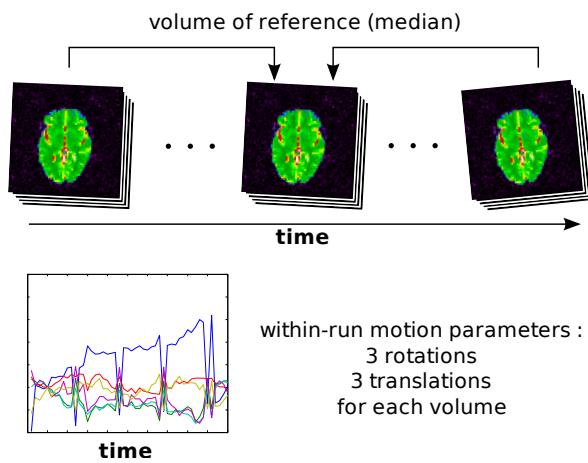


Figure 1.4: Motion estimation based on rigid-body motion estimation of the functional volumes, the procedure provide 6 motion parameters for each volume (3 translation and 3 rotation) Schematic of the preprocessing pipeline including spatial and functional normalization (from NIAK preprocessing pipeline ⁴).

1.4 Multi-site

In most experiments conducted in neuroimaging, the main factors that influence power are: (1) the size of the effect, determined by the difference of the mean connectivity of one group versus a control group and the variability of this difference across subjects and groups; (2) the probability of rejecting the null hypothesis when it is true; and (3) the sample size, i.e. the number of subjects in the study (Desmond and Glover 2002). This last factor is usually the only one controlled by the investigator, hence why an increasing number of researchers share multicentric, sometimes multiprotocol,

data suitable to statistical analysis. In research it is very difficult to obtain a grant large enough to scan a cohort larger than 80 subjects, therefore researcher and consortium initiatives have started to pool their resources together to make initiative composed of publicly available large cohorts of subjects like the 1000 functional connectome (Biswal et al. 2010), ADNI (Mueller et al. 2005), among others. In clinical trial the justification for multicentric acquisition is more of a logistical one then a financial reason; they need to recruit a large amount of subject in a short period of time. In order to achieve this goal they mandate the recruitment to multiple clinical centers across the globe which accelerate the evaluation time of a drug. Although these centers may be similar by their scanner protocols, scanners will have difference in their software version, specific add-on to the scanners, and, most importantly, vendors (even field strength may differ in some cases). Unfortunately between studies, MR acquisition methodologies are among the most commonly cited sources of measurement variation (Friedman et al. 2006). This is why it is important to assess if multi-site resting-state connectivity analysis are feasible (we can combine the data from multiple sources while introducing a reasonable amount of variance which is still acceptable to detect effects in the data) and what corrective measure on the data should be applied to reduce the bias introduced by multi-site analysis. Among the factor of variability across sites, we can list the following 3 categories described in (Yan et al. 2013b):

1. Acquisition-related variations:

- (a) *Scanner make and model (Friedman et al. 2006)*
- (b) *Sequence type (spiral vs. echo planar; single-echo vs. multi-echo) (Klarhofer et al. 2002), parallel vs. conventional acquisition (Feinberg et al. 2010) (Lin et al. 2005)*
- (c) *Coil type (surface vs. volume, number of channels, orientation).*
- (d) *Acquisition parameters: repetition time, number of repetitions, flip angle, echo time, and acquisition volume (field of view, voxel size, slice thickness/gaps, slice prescription) (Friedman and Glover 2006).*

2. Experimental-related variations:

- (a) *Participant instructions (Hartstra et al. 2011), eyes-open/eyes-closed (Yan et al. 2009) (Yang et al. 2007), visual displays, experiment duration (Fang et al. 2007) (Van Dijk et al. 2010).*

3. Environment-related variations:

- (a) *Sound attenuation measures (Cho et al. 1998) (Elliott et al. 1999).*
- (b) *Attempts to improve participant comfort during scans (e.g., music, videos) (Cullen et al. 2009).*
- (c) *Head-motion restraint techniques (e.g., vacuum pad, foam pad, bite-bar, plaster cast head holder) (Edward et al. 2000) (Menon et al. 1997).*
- (d) *Room temperature and moisture (Vanhoutte et al. 2006).*

In 2009, the publicly released 1000 Functional Connectomes Project (FCP) and International Neuroimaging Data-sharing Initiative (INDI) provided a glimpse of the variability in imaging methodologies employed by the neuroimaging field. The dataset includes rs-fMRI samples independently collected at imaging sites around the world. A noteworthy aspect of this dataset is the variation in almost every parameter of the imaging acquisition methodologies, while the majority of subject-related variables are not reported (due in most cases, to the fact that they were not thoroughly recorded). Despite justifiable scepticism, feasibility analyses demonstrated that meaningful explorations of the aggregate dataset, composed of 24 imaging sites for a grand total of 1093 subjects, could be performed (Biswal et al. 2010). Although no explicit correction for multi-site variability was used, they only use global signal correction (GSC) to normalize subjects which may introduce anti-correlation in the data (Carbonell et al. 2014, Fox et al. 2009, Murphy et al. 2009, Power et al. 2014, Saad et al. 2012). After accounting for site-related differences, the analysis showed brain-behaviour relationships with phenotypic variables such as age, gender, and diagnostic label, and confirmed a variety of prior hypotheses (Biswal et al. 2010, Fair et al. 2012, Tomasi and Volkow 2010, Zuo et al. 2012). While

encouraging, many uncontrolled and unknown factors in the 1000 FCP remain a source of concern, as they spread beyond simple site effects and can limit the datasets utility as highlighted by Yan et al. (2013a). An other compelling proof of multi-site bias is the study reported by Nielsen et al. (2013) where they did an analysis on a single site dataset and a multi-site dataset of subject with autism and concluded that the multi-site autism study classification accuracy significantly outperformed chance but was much lower for multi-site prediction than for previous single site results (Nielsen et al. 2013). We therefore need to keep in mind that the site effect must be taken in account in the analysis or we may reduce our detection power.

1.5 Resting-state connectivity

Resting-state (RS) functional connectivity captures slow fluctuations of hemodynamic activity without performing any task. These temporal fluctuations can be monitored using the signal measured with fMRI. The first study that introduced the concept of resting state functional connectivity was the one of Biswal et al. (1995). By performing a task activation response to bilateral left and right finger tapping, Biswal and colleagues were able to activate the corresponding motor areas on the cortex. In a second analysis, they considered one of the area activated during the task as a seed region for a seed-based analysis in resting-state condition. Seed-based analysis consists in detecting temporal correlation between the signal of the predefined seed area and the time course of all the voxels of the brain. Using this seed-based analysis, they found RS correlations between similar brain regions than the ones involved during the task. A more recent review done by Fox and Raichle (2007) illustrated in Figure 1.5 show the ability to identify the complete sensorimotor network using only the bold signal from a small region in that network.

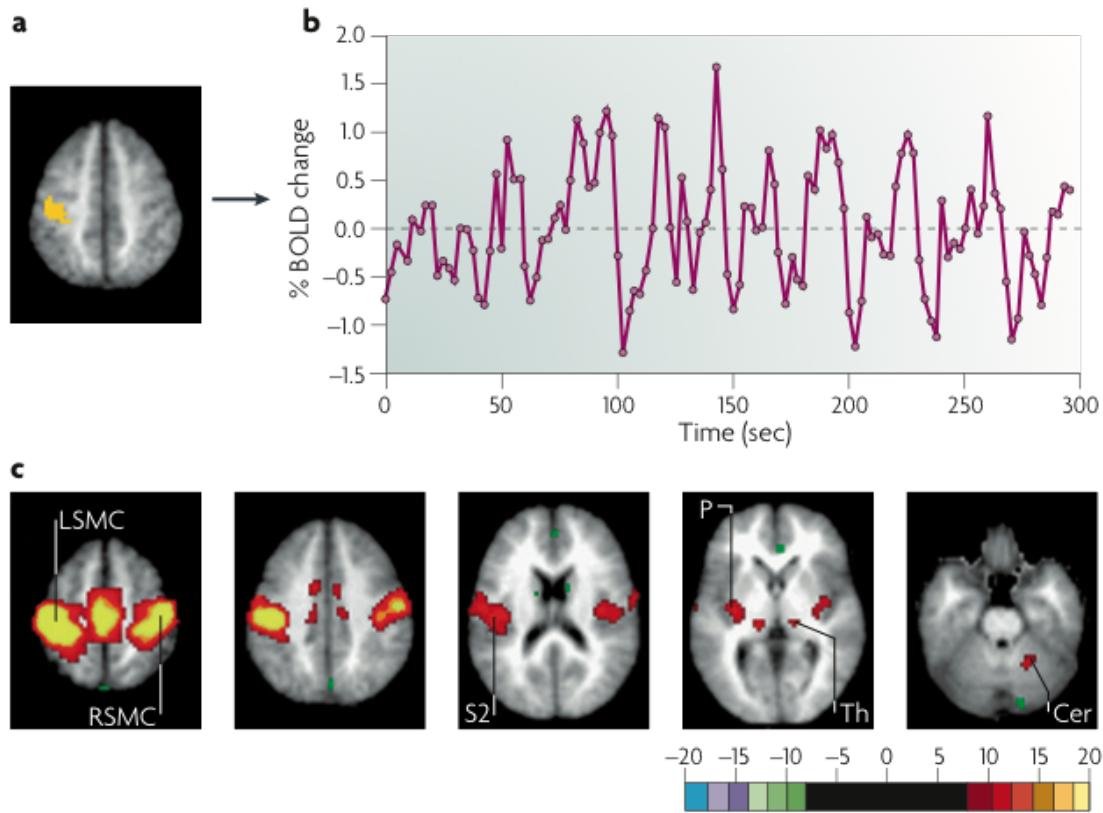


Figure 1.5: Generation of resting-state correlation maps. a) Seed region in the left somatomotor cortex (LSMC) is shown in yellow. b) Time course of spontaneous blood oxygen level dependent (BOLD) activity recorded during resting fixation and extracted from the seed region. c) Statistical z-score map showing voxels that are significantly correlated with the extracted time course. Their significance was assessed using a random effects analysis across a population of ten subjects. In addition to correlations with the right somatomotor cortex (RSMC) and medial motor areas, correlations are observed with the secondary somatosensory association cortex (S2), the posterior nuclei of the thalamus (Th), putamen (P) and cerebellum (Cer) (Fox and Raichle 2007).

In resting-state acquisition, no task is performed during the scan; the subject is instructed to rest with his eyes open or closed, as opposed to the task-based acquisition where the subject has to perform a specific task. These early results from Biswal et al.

suggest that it is possible to identify the functional organization of different structures without doing any specific task, just by looking at spontaneous fluctuations in brain activity. Several studies have demonstrated that patterns extracted from temporal correlations of rs-signals within the brain volume are organized in space and have a good reproducibility from subject to subject: the so-called consistent resting state networks (CRSN) (Damoiseaux et al. 2006). Each network is a combination of multiple brain regions or units, not necessarily spatially close to each other, which are sharing similar low frequency fluctuations of the BOLD signal; this is usually represented as a functional connectivity matrix where one column of the matrix represent the connectivity of a region with the rest of the brain called a functional connectivity map (see Figure 1.6 for a graphical representation of the two concepts).

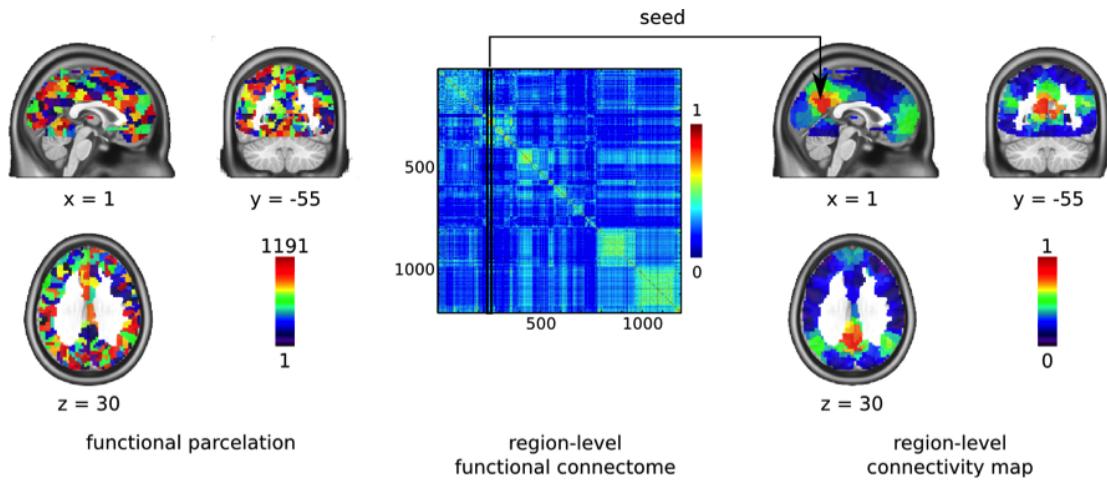


Figure 1.6: Functional connectome: on the left a representation of a functional parcelation, in the middle a region-level functional connectome representing the connectivity between each pair of region and on the right the connectivity map based on a region of interest extracted from the functional connectome.

Several techniques have been used to identify these so-called resting-state networks (RSN). These networks show the functional organisation of various brain regions and several studies have demonstrated the relation of those functional networks to specific tasks in humans as well as in animal models (Biswal et al. 1995) (Buckner et al. 2008)

(Greicius et al. 2009). Cordes et al. describe that several functional networks (sensorimotor, language and visual) identified in RS condition were exhibiting similar regions than the ones involved when performing a specific task (motor, language, or visual task) (Cordes et al. 2000). Based on such observations, CRSNs were then labeled according to the corresponding brain areas found active when performing such a specific task (e.g. motor or cognitive for example) thus associated with a brain function: e.g. auditory, visual, language networks (Cordes et al. 2000) see Figure 1.7 for a list of common RSN. It enables us to establish the relationship between regions active when performing a specific task and RSNs, thus supporting the functional relevance of these networks.

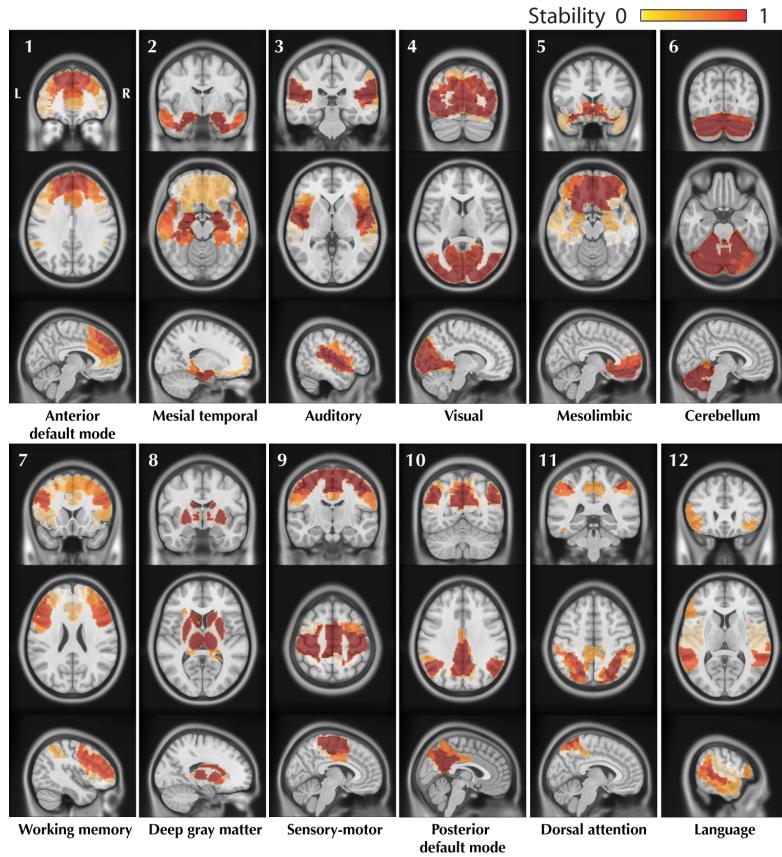


Figure 1.7: The figure shows 12 CRSNs identified using BASC (Bootstrap Analysis of Stable Cluster (Bellec et al. 2010b)) group level analysis of 25 healthy control subjects. BASC is a clustering based method using evidence accumulation for the identification of stable cluster. For each CRSN: 3 slices (coronal, axial, sagittal) are shown superimposed on an anatomical MRI template (MNI152). Labelling of each network was done visually based on previously reported CRSNs in the literature. The figure shows the usual networks: Default Mode Network (#1,#10), Auditory (#3), Visual (#4), Sensory-Motor (#9), Attention (#7,#11) and Language(#12). BASC also identified 4 other networks, less often reported, but characterized by high statistical stability: Mesio-Temporal (#2), Mesolimbic (#5), Cerebellum (#6) and Deep Gray Matter (#8) (Dansereau et al. 2014a) in press.

Subsequently, rs-fMRI signal have been used in healthy subjects to investigate nor-

mal brain function, within various functional systems, such as auditory (Cordes et al. 2001), visual (Lowe et al. 1998), language (Hampson et al. 2002), limbic systems (Greicius et al. 2003, Tian et al. 2007, Wink and Roerdink 2006) and motor (Jiang et al. 2004, Lowe et al. 1998). Interestingly, resting-state fMRI signals have also been used to characterize the pathophysiological changes of some diseases, such as multiple sclerosis (Lowe et al. 2002), epilepsy (Waites et al. 2006), schizophrenia (Liang et al. 2006, Salvador et al. 2007, Zhou et al. 2007; 2008), attention deficit hyperactivity disorder (Tian et al. 2007, Zang et al. 2007), blindness (Liu et al. 2007, Yu et al. 2007), major depression (Anand et al. 2005, Greicius et al. 2007) and acute brainstem ischemia (Salvador et al. 2005). Thus, we believe that resting-state fMRI will be an increasingly important modality for exploring the functional abnormalities of patients with AD. Since a clinically meaningful hierarchical structure exist in the brain it is possible to use this structure to reduce the dimensionality of the problem and variability of the data while providing an informed organisation of the structure to a machine learning procedure looking to find functional markers characteristic of specific pathological populations (e.g. Alzheimer's disease).

1.6 Prediction

1.6.1 Prediction in the context of AD

In the past few years, several major studies have been initiated that have aimed to predict who will develop AD at the prodromal or even asymptomatic stage, with the ultimate goal of providing a platform for therapeutic intervention with disease-modifying therapies. Many of these studies were designed to evaluate the role of neuroimaging and chemical biomarkers in assessing and predicting progression in individuals without cognitive impairment and in individuals with MCI. A very small literature currently reports findings on AD using rs-fMRI and they usually have outstanding performance like (Jiang et al. 2014) (94%). Unfortunately, the authors of this study performed a 10-fold cross-validation which did not include the feature selection process, which potentially lead to overestimation of the real accuracy. The rest of published studies have mainly

used leave one out cross-validation and small sample size, which raises some concern with regards to the generalization ability of those trained predictors, e.g. (Chen et al. 2011) (87%), (Dai et al. 2014) (80%)).

1.6.2 Importance of preprocessing to improve classification

Preprocessing and normalization of the source data is a crucial point that may affect greatly the resulting classification or potential outcome measures. Moreover the introduction of multicentric data in the classification will render the task even more difficult, Nevertheless, a multi-site cohort helps test generalizability of the results across different samples, making it more likely that connections identified as predictive of a disease state indeed reflect generic traits of the pathology rather than particularities of a single dataset.

1.6.3 High dimensionality problem

In rs-fMRI, we obtain signal from brain activity at the voxel level representing more than 10^4 voxels in the gray matter cortex, where the vast majority of the neurons are located. Although preprocessed, these data continue to have a lot of variance and we need to identify functional organisation more meaningful in term of clinical interpretation as well as an improved representation of the feature space that enhance the characterization of the functional modules. We can commonly extract functionally and clinically meaningful network features using principal component analysis (PCA) (Zhong et al. 2009), independent component analysis (ICA) (McKeown et al. 1998), various clustering algorithms such as k-means (Baumgartner et al. 1998), hierarchical clustering (Cordes et al. 2002), normalized cut-graph (van den Heuvel et al. 2008), self-organizing maps and neural gas (Meyer-Baese et al. 2004). Once we have functional parcels of the brain it is important to find the right metric to evaluate the connectivity between regions of the brain at the individual level. The most common metric used in the field of connectivity is the Pearson's correlation coefficient between the time series associated with a pair of regions.

1.6.4 Feature selection

Feature selection is often a critical step prior to any learning algorithm. The selection reduces the computational complexity of learning algorithms and exposes potentially clinically meaningful information. In many cases, this process can also improve the prediction accuracy by removing redundant and irrelevant information. Therefore feature selection is an important step in effective learning of large data sets. The features selection methods are usually organized into two categories: filter methods and wrapper methods (Kotsiantis 2007). The filter method evaluates the relevance of features by looking only at the properties of the data without any knowledge on the classification labels. The wrapper method assesses the goodness a feature subset using the performance of a learning algorithm. The lack of reproducibility of reported markers (subset of features) is one of the main obstacle for the adoption of such marker in a clinical setup. If the marker is indeed discriminative of the pathology or of its progression we would expect the same features to be selected and exhibit similar performance across various studies.

1.6.5 Classification

Machine learning methods have become very popular to classify functional brain images (Costafreda et al. 2009, Fu et al. 2008, Hahn et al. 2011, Marquand et al. 2008, Nouretdinov et al. 2011), for example to discriminate them between healthy and diseased populations. The most popular machine learning techniques in fMRI arguably is support vector machine (SVM) (Cortes and Vapnik 1995) which has been used in the past to categorize individual structural or functional brain images by differentiation of images from two groups (e.g. patient/control or male/female) (Fan et al. 2005, Kawasaki et al. 2007, Lao et al. 2004, Mourao-Miranda et al. 2005). A second widely used method is linear discriminant analysis (LDA) classifier; the main advantage of LDA is the ability to easily include confounding effects in the decision model, such as the inter-site difference in average connectivity measures.

1.7 Objectives

The objectives of the theses is to 1) reduce motion-related bias in connectivity measures, 2) evaluate the feasibility of multicentric fMRI analysis as well as identifying normalization procedures to reduce the between-site variance, and 3) design a prediction pipeline for the data-driven identification of biomarkers of AD in resting-state fMRI. The pipeline will have a feature selection tool to find a highly reduced set of functional connections with good prediction accuracy for the future conversion to a dementia of the Alzheimer's type in individuals with mild cognitive impairment.

CHAPTER 2

SCRUBBING AND MOTION ARTEFACT CORRECTION

2.1 Introduction

Head motion is probably the most severe problem in fMRI studies, although it cannot be avoided. The quality of the fMRI data is strongly affected by the presence of head motion. As previously mentioned, a number of publications have proposed post-hoc correction in order to reduce head motion artefact. Although head motion can be corrected in the image space, displacement of the head reduces the homogeneity of the magnetic field, which is fine-tuned prior to functional scans for a given head position. Since head displacements lead to non-optimal tuning, motion artefacts are not fully removed even after perfect realignment of successive functional volumes in image space. The impact of motion and the method available to reduce its effect on functional connectivity thus need to be carefully evaluated. Typically subjects with average motion over 3 mm are excluded from the analysis. A first popular technique to reduce motion artifact is global signal correction (GSC), which consists of regressing the average of all the time-series in the brain. This technique has raised a number of concerns due to the potential artificial introduction of anti-correlation in the data (Carbonell et al. 2014, Fox et al. 2009, Murphy et al. 2009, Power et al. 2014, Saad et al. 2012). Another corrective measure is the CompCor method (Behzadi et al. 2007) that regresses out the n first components of a principal component analysis (PCA) on the time-series of the white matter and cerebrospinal fluids voxels. As no signal of neuronal origin are expected in such areas, the CompCor method was suggested to provide a compact representation of the structured noise present in an fMRI dataset. Both GSC and CompCor are global corrections, which are not specific to motion correction. The GSC in particular incorporates signal from the grey matter regions. Power and colleagues showed in 2012 that spurious but systematic correlations in fc-MRI arise from subject motion even at levels previously regarded as insignificant (variation smaller than 1 mm / degree in motion parameters) and are

not adequately removed by common functional connectivity processing steps. Power et al. (2012) suggested using a method called “scrubbing” (or censoring) in order to reduce motion-related bias from fMRI time-series. With the scrubbing method, the frames with motion level above a certain threshold are simply removed from the time-series. The metric used to quantify the level of motion from one frame to the other is called framewise displacement (FD). It is calculated as the sum of the absolute values of the differentiated realignment parameters estimated at every time point, giving an approximate conversion of rotation parameters ($\alpha_i, \beta_i, \gamma_i$) in a unit equivalent to a translation in millimeters on a sphere of radius 50mm and translation parameters (d_{ix}, d_{iy}, d_{iz}) for the other rigid body parameters. The main limitation of the scrubbing method is the loss of time-points, which can lead to subjects loss due to an insufficient number of time point remaining after scrubbing correction. The question is therefore to evaluate if the trade-off between data loss and quality is detrimental or beneficial to statistical power. I have recently published results concerning this issue Dansereau et al. (2014b) (published proceeding and manuscript in preparation). In this work we aimed to (1) characterize the impact of motion on rs-fMRI in cognitively normal elderly (CNE) participants as well as patients with mild cognitive impairment (pMCI) and dementia of the Alzheimer’s type 0.5-No (pDAT), and, (2) evaluate how the scrubbing impacts the differences in connectivity between groups (CNE, pMCI, pDAT).

$$\Delta d_{ix} = d_{(i-1)x} - d_{ix}, \text{ derivative of a translation parameter in the x axis} \quad (2.1)$$

$$FD_i = |\Delta d_{ix}| + |\Delta d_{iy}| + |\Delta d_{iz}| + |\Delta \alpha_i| + |\Delta \beta_i| + |\Delta \gamma_i|. \quad (2.2)$$

2.2 Method

2.2.1 Dataset

This project was based on a real dataset with fMRI data acquired in different settings. The dataset included 313 elderly adults aggregated across 5 independent studies: the ADNI2 study and 4 other studies based in Montreal, Canada. The dataset included three different clinical cohorts, for a grand total of 126 CNE participants (51 males, age range = 57-94 yrs), 133 patients with MCI (70 males, age range = 55-89 yrs), and 54 patients with DAT (22 males, age range = 55-88 yrs). We also included 355 cognitively normal young adults (CNY) from the 1000 functional connectome project (150 males, age range = 18-46 yrs) as a reference dataset (Biswal et al. 2010). This project was approved by the local ethics review board.

2.2.2 Preprocessing

The datasets were analysed using the NeuroImaging Analysis Kit (NIAK¹) version 0.12.14, under CentOS version 6.3 with Octave² version 3.8.1 and the Minc toolkit³ version 0.3.18. Analyses were executed in parallel on the "Mammouth" supercomputer⁴, using the pipeline system for Octave and Matlab (Bellec et al. 2012), version 1.0.2. Brain map visualizations were created using MRICron software (Rorden et al. 2007). Each fMRI dataset was corrected of inter-slice difference in acquisition time and the parameters of a rigid-body motion was estimated for each time frame. Rigid-body motion was estimated within as well as between runs, using the median volume of the first run as a target. The median volume of one selected fMRI run for each subject was coregistered with a T1 individual scan using Minctracc (Collins et al. 1998), which was itself non-linearly transformed to the Montreal Neurological Institute (MNI) template (Fonov

¹<http://www.nitrc.org/projects/niak/>

²<http://gnu.octave.org>

³<http://www.bic.mni.mcgill.ca/ServicesSoftware/ServicesSoftwareMincToolKit>

⁴<http://www.calculquebec.ca/index.php/en/resources/compute-servers/mammouth-parallel-e-ii>

et al. 2011) using the CIVET pipeline (Zijdenbos et al. 2002). The MNI symmetric template was generated from the ICBM152 sample of 152 young adults, after 40 iterations of non-linear coregistration. The rigid-body transform, fMRI-to-T1 transform and T1-to-stereotaxic transform were all combined, and the functional volumes were resampled in the MNI space at a 3 mm isotropic resolution. The 'scrubbing' method of (Power et al. 2012), was used to remove the volumes with excessive motion with three cut-offs points: no scrubbing ($FD \geq 0$), $FD \geq 0.5$ and $FD \geq 0.2$. A minimum number of 50 unscrubbed volumes per run, corresponding to ~ 125 s of acquisition for a TR of 2.5 seconds, was then required for further analysis. For this reason, some subjects were rejected from the subsequent analyses: 11 CNE, 13 pMCI and 3 pDAT for a scrubbing at $FD \geq 0.5$ and 83 CNE 95 pMCI and 39 pDAT for a scrubbing at $FD \geq 0.2$ (see table 2.I). The following nuisance parameters were regressed out from the time series at each voxel: slow time drifts (basis of discrete cosines with a 0.01 Hz high-pass cut-off), average signals in conservative masks of the white matter and the lateral ventricles as well as the first principal components (95% energy) of the six rigid-body motion parameters and their squares (Lund et al. 2006),(Giove et al. 2009). The fMRI volumes were finally spatially smoothed with a 6 mm isotropic Gaussian blurring kernel.

2.2.3 Statistical analysis

Impact of motion and scrubbing level on connectivity was estimated with a *t*-test and a group false-discovery rate (FDR) (Hu et al. 2010). Group differences (CNE, pMCI and pDAT) were investigated using *t*-tests for ten point-to-point connections selected based on a literature review (Dansereau et al. 2013) including covariates to model age, gender and site-specific bias using model averaging (Willer et al. 2010). The significance of the difference in average connectivity between the clinical cohorts was assessed by a *t*-test in a linear model, including covariates to model site-specific bias. This procedure lead to a *p*-value estimating the false positive rate when testing against the null hypothesis of no group difference. To test the robustness of the group differences, the *t*-tests ($p \leq 0.05$) were replicated $B = 10^4$ times using random subsamples including 70% of each group. For each replication b , a *p*-value p^{*b} was generated. The stability of detection of a group

difference as significant was estimated by the following Monte-Carlo average:

$$B^{-1} \sum_{b=1}^B \left(p^{*b} \leq 0.05 \right). \quad (2.3)$$

2.3 Results

2.3.1 Motion distribution in different clinical cohorts

A wide range of motion level was observed in the sample (Figure 2.1) and over all the elderly population seems to have more motion than the CNY. Even when scrubbing aggressively the difference in FD distribution between groups remained. We therefore needed to account for FD differences in subsequent analysis. In Figure 2.1 a good overlap in the FD distribution across groups has been observed, the effect of scrubbing reduce the spread of the distribution and center the average FD around 0.2 (for scrubbing with a threshold $FD \geq 0.5$) and 0.1 (for a threshold of $FD \geq 0.2$). Overall the amount of motion observed in CNY subjects was smaller than the one observed in the elderly populations (ECN, pMCI, pDAT) as shown in the boxplot representation on the left of Figure 2.1. In all preprocessing strategies the CNY remain significantly different in terms of FD compared to all the elderly groups. In term of the comparisons in FD distribution between elderly cohorts, scrubbing was able to remove some of the variability between groups but not sufficiently to remove the significant differences in FD in the CNE minus pMCI and CNE minus pDAT comparisons. The differences between pMCI and pDAT were not significant regardless of the preprocessing strategy. Surprisingly, results also showed that in the elderly population, the pDAT population was the one with the smallest amount of motion, followed by pMCI and finally CNE. There was no indication that the motion artifacts (and corresponding data quality) would be inferior in clinical populations. The is a significant difference in FD distribution between CNY and all the other groups in the 3 preprocessing strategies (see Figure 2.1 on the right). Significant differences in term of FD distribution arose when scrubbing at $FD \geq 0.5$ and $FD \geq 0.2$ between the pDAT group and pMCI as well as pDAT and CNE group.

Regarding the impact of scrubbing on the retention of the original dataset, 90% of

subjects survived the exclusion criteria of a minimum of 50 frames after scrubbing at $FD \geq 0.5$ and 30% survived an aggressive scrubbing of $FD \geq 0.2$, see Table 2.I for the specific breakdown by clinical cohorts of elderly subjects (CNE, pMCI and pDAT). Only 1% of the subjects were lost in the CNY cohort (99% retention) for an $FD \geq 0.5$ and 65% survived for $FD \geq 0.2$.

	CNE	pMCI	pDAT	Total	CNY
No threshold	126	133	54	313	355
$FD \leq 0.5$	115 (91%)	120 (90%)	51 (94%)	286 (91%)	352 (99%)
$FD \leq 0.2$	43 (34%)	38 (29%)	15 (28%)	96 (31%)	230 (65%)

Table 2.I: Retention rate for CNY, CNE, pMCI and pDAT at various scrubbing levels (standard, scrubbing $FD > 0.5$ and scrubbing $FD > 0.2$).

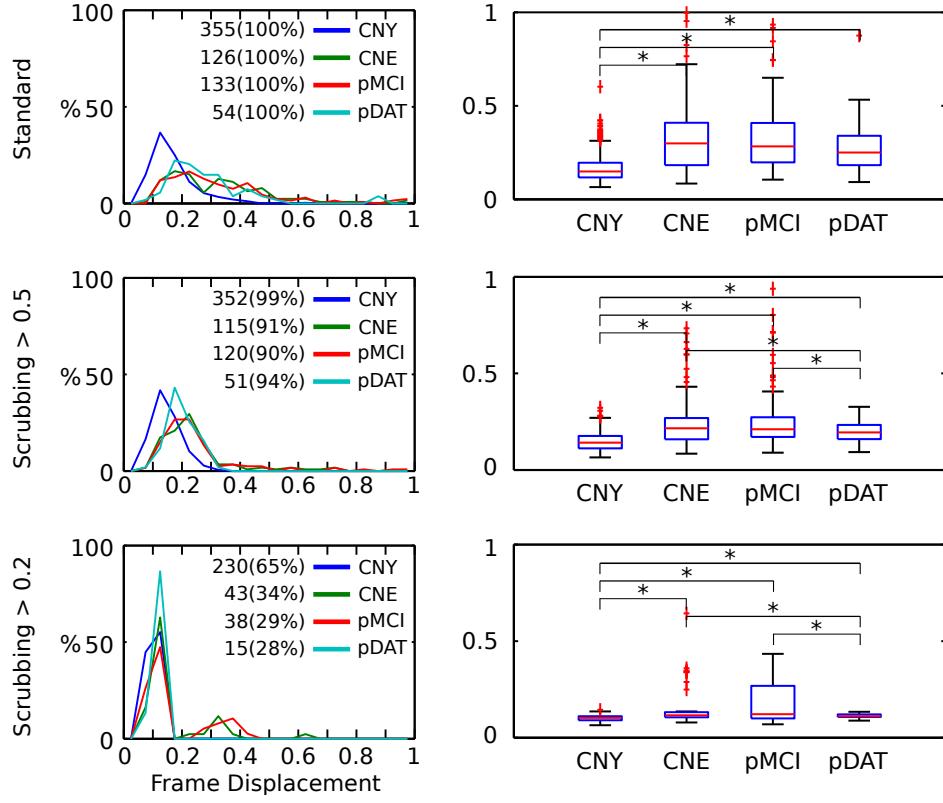


Figure 2.1: Distribution of the frame displacement (FD) for 3 groups (CNE, pMCI, pDAT) when scrubbing was applied at various levels (no scrubbing, scrubbing of $FD > 0.5$ and scrubbing of $FD > 0.2$). The boxplot on the right showed the distribution of FD with their associated statistical differences t -test (marked with a * for a $p < 0.05$).

2.3.2 Default mode network in young adult and elderly population

The default-mode network (DMN) is hypothesized to be a key target of neurodegeneration in Alzheimer's disease. A visual inspection of the DMN using a standard preprocessing strategy showed more positive correlation values in the CNE population compared to the CNY (See Figure 2.2), probably reflecting increased partial-volume effect with cerebro-spinal fluids due to the large amount of cortical atrophy typically seen in aging. Moreover, a slight decrease in connectivity was observed in the frontal part of the DMN in the CNE population compared to the more common pattern of fronto

parietal connectivity shown in the CNY population. This finding of more negative correlation for CNE subjects compared to CNY subjects was consistently observed regardless of the preprocessing strategies except for GSC, see Figure 2.2. The GSC seemed to increase the extent of the negative correlation found in other preprocessing strategies, as previously reported in the literature, e.g. (Murphy et al. 2009).

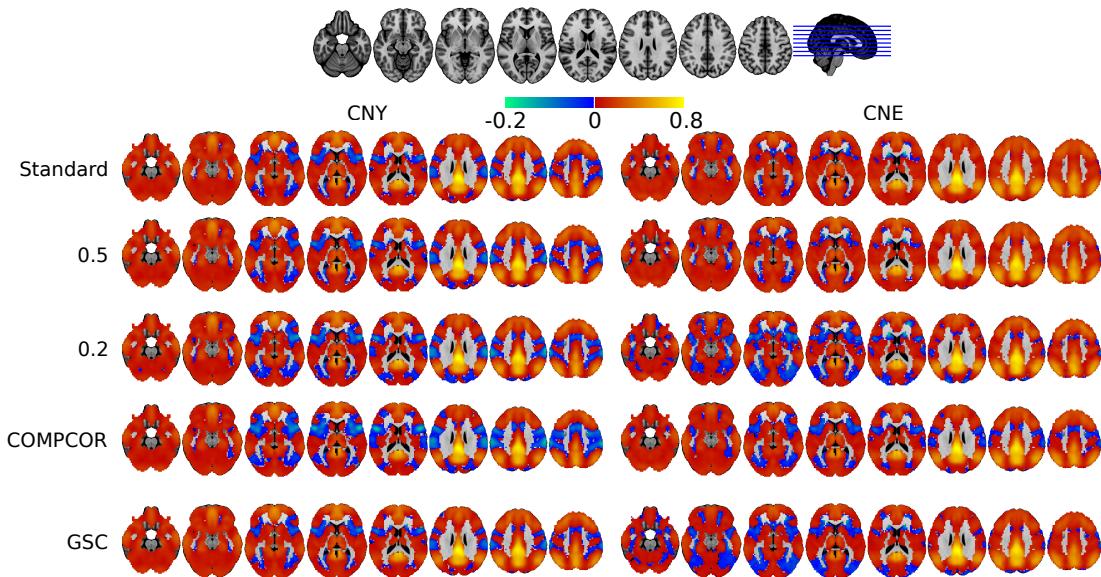


Figure 2.2: Overlay of the average default mode network (DMN) with various preprocessing strategies for CNY and CNE populations on the ICBM 152 anatomical atlas. Seed based maps with a seed in the PCC using Fisher z transform of the Pearson's r correlation between the average time series of each network.

The connectivity patterns affected by scrubbing are consistent across various groups with and without dementia see Figure 2.3 for an example of three groups. The scrubbing procedure significantly increased connectivity strength inside the default-mode network and reduced connectivity with anti-correlated regions (Figure 2.3) consistently across all groups. More global decrease in connectivity changes can be observed for the CompCor and GSC methods. For CompCor a strong change in the sensory motor network can be observed as well as an increased connectivity in the ventral part mesio-frontal cortex.

For the GSC method decreases in connectivity can be observed across the brain as well as a strong decrease in the occipital lobe.

2.3.3 Impact of scrubbing on the connectivity

2.3.3.0.1 Global impact of scrubbing Scrubbing significantly increased connectivity strength inside the default-mode network and reduced connectivity with anti correlated regions. Difference in functional connectivity of the DMN show significant increase of the frontal part of the DMN with scrubbing at $FD \geq 0.5$ and $FD \geq 0.2$ a region normally positively correlated with the PCC (see 2.3). Significant decrease of connectivity with region associated with attention are also observed (dorsal attention network). The preprocessing with CompCor show massive decrease in the sensorimotor region and globally across the brain except for a mesio frontal region (anterior cingulate cortex ACC) more ventral than the expected mesio frontal region associated with the DMN (medial prefrontal cortex MPFC). The map of difference for the GSC show massive decrease in connectivity across the brain and in particular the occipital lobe, premotor and sensorimotor areas. The effect reported for the preprocessing with GSC are even more pronounced for the CNY group (see supplementary material Figure 2.4).

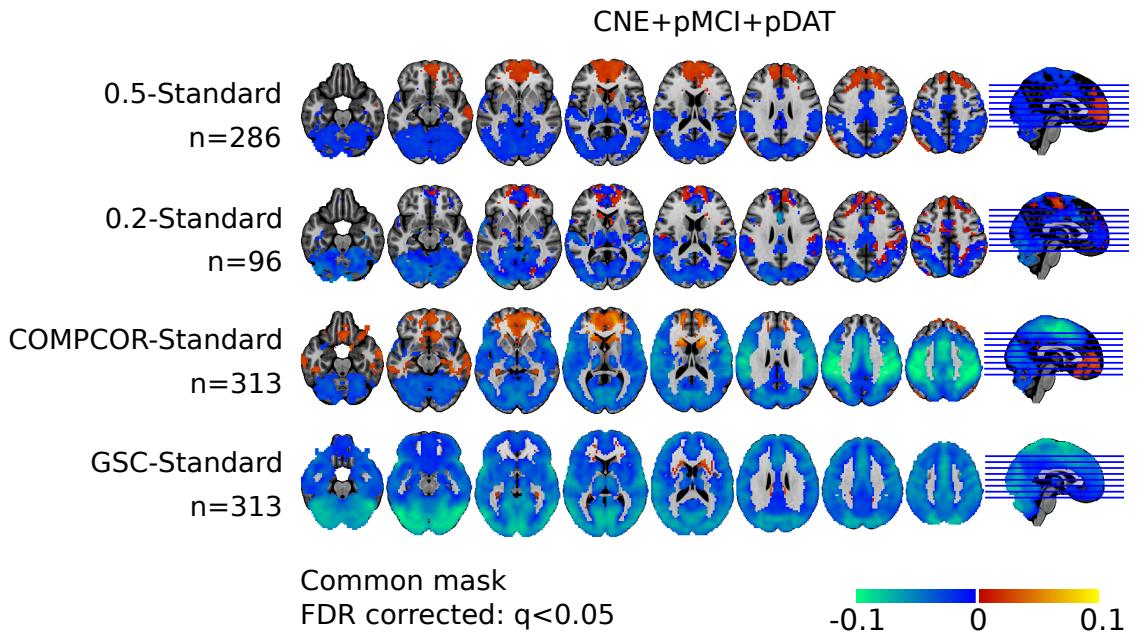


Figure 2.3: Differences in functional connectivity for the default mode network (seed in the PCC). Differences in connectivity between all the groups pooled together (CNE, pMCI, pDAT) with scrubbing ($FD > 0.5$ and $FD > 0.2$), CompCor, and GSC. The mask used depicts only significant result of the t -test (FDR correction $q < 0.05$) only the two scrubbing procedures use the union of their respective mask (common mask).

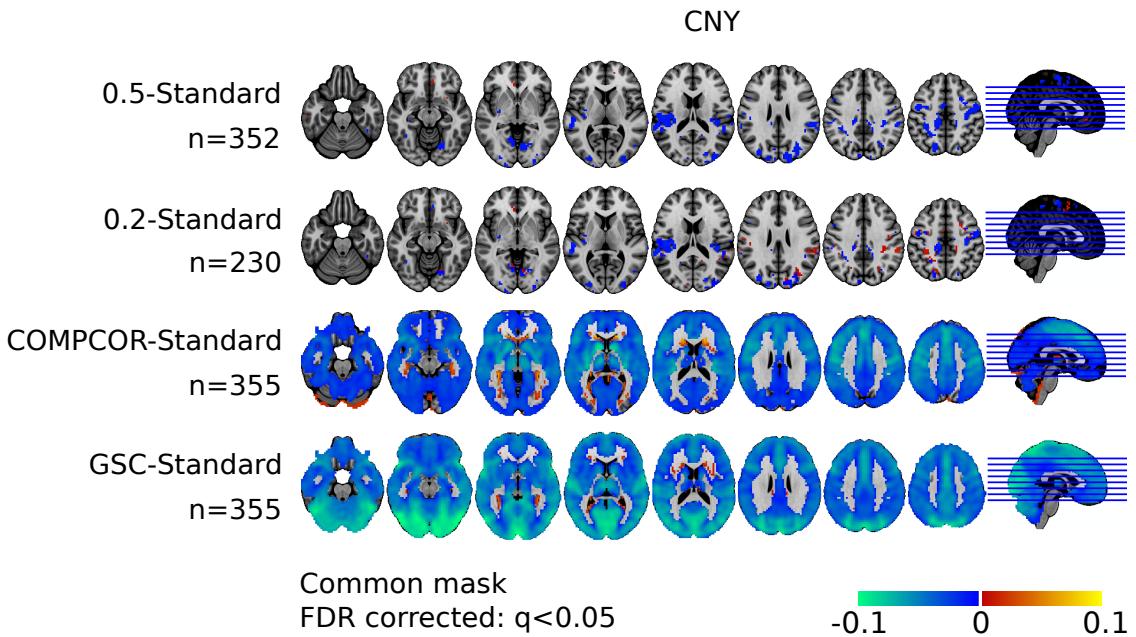


Figure 2.4: Differences in functional connectivity for the default mode network (seed in the PCC). Differences in connectivity between all the CNY with scrubbing ($FD > 0.5$ and $FD > 0.2$), CompCor, and GSC. The mask used depict only significant result of the t -test (FDR correction $q < 0.05$) only the two scrubbing procedures use the union of there respective mask (common mask).

2.3.3.0.2 Population specific impact of scrubbing The population specific difference in connectivity (see Figure 2.5) show almost identical findings as the combination of CNE, pMCI and pDAT therefore confirming that the observation are transferable in every population and not specific. Although the increase connectivity of the MPFC is observed in all groups when scrubbing is applied, the difference in connectivity is greater as we progress toward dementia (difference in fc for MPFC area CNE < pMCI < pDAT). Difference map for CompCor and GSC revealed an invers patern where the greatest changes are observed in the CNE group followed by pMCI and finally pDAT for negative differences. CompCor show a increase in connectivity for the ACC area in the three groups but particularly in the pMCI group followed by CNE and then pDAT.

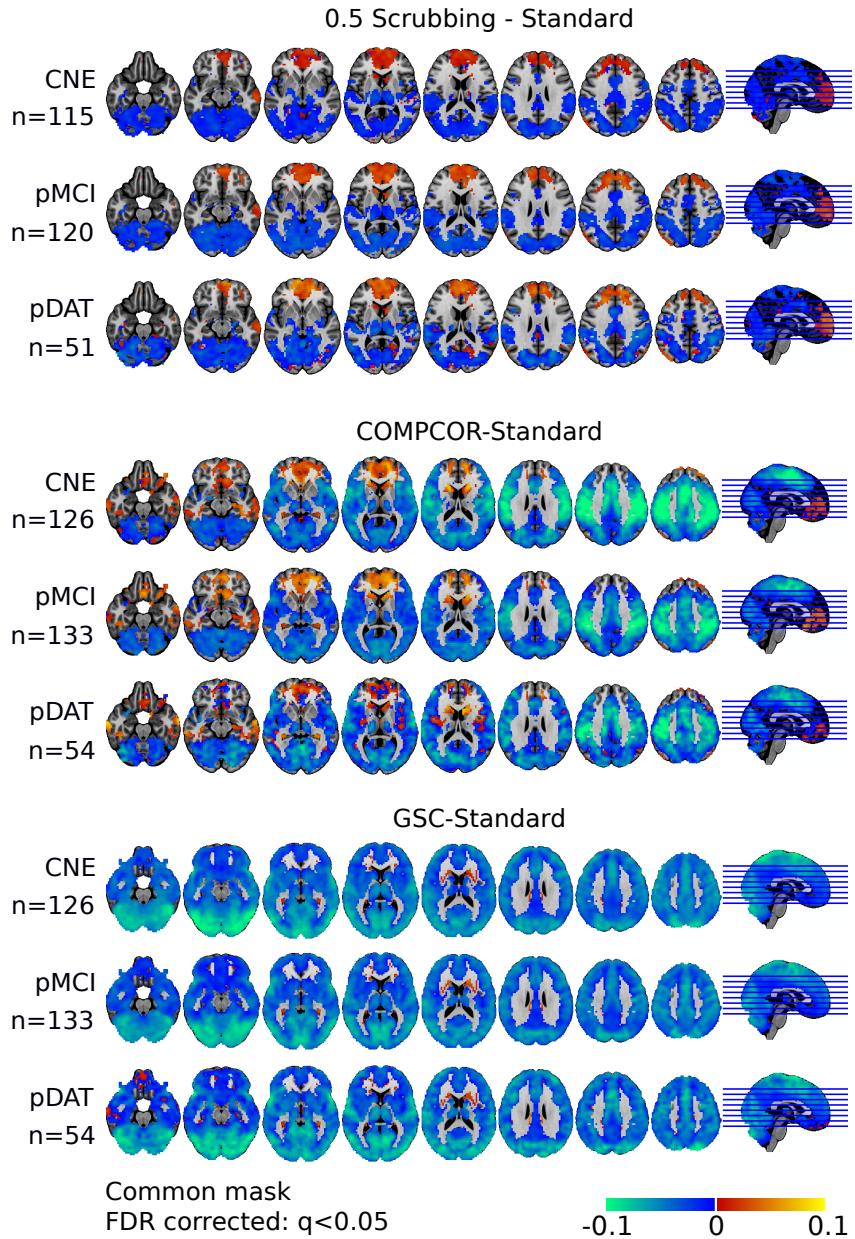


Figure 2.5: Differences in functional connectivity for the default mode network (seed in the PCC). Differences in connectivity for each group (CNE, pMCI, pDAT) compared to baseline (standard preprocessing) with and without scrubbing ($FD > 0.5$), CompCor, and GSC (FDR correction $q < 0.05$) for all voxels showing a significant effect in at least one of the contrast.

In order to assess if the significant differences in connectivity between groups are affected by the preprocessing strategy we did a t -test on each preprocessing strategy for every contrast (pDAT-CNE 2.6, pDAT-pMCI 2.7 and pMCI-CNE 2.7).

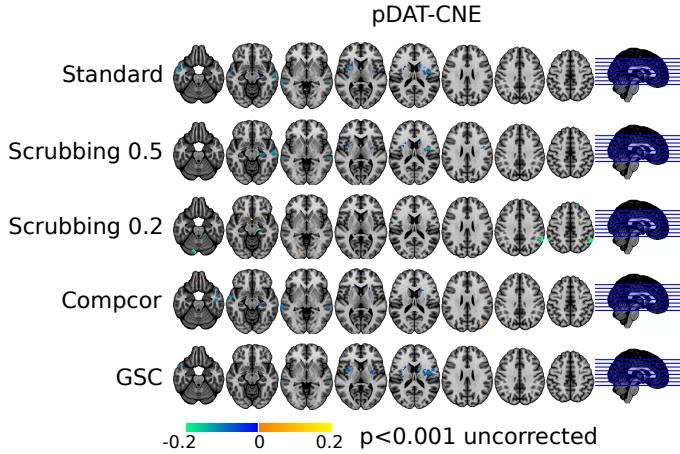


Figure 2.6: Impact of preprocessing on connectivity differences between the pDAT and CNE groups. Connectivity differences between DMN (t -test $p < 0.001$ uncorrected) computed with various preprocessing strategies (standard, scrubbing (0.5 and 0.2), CompCor and GSC). The maps are represented on top of the ICBM 152 anatomical atlas.

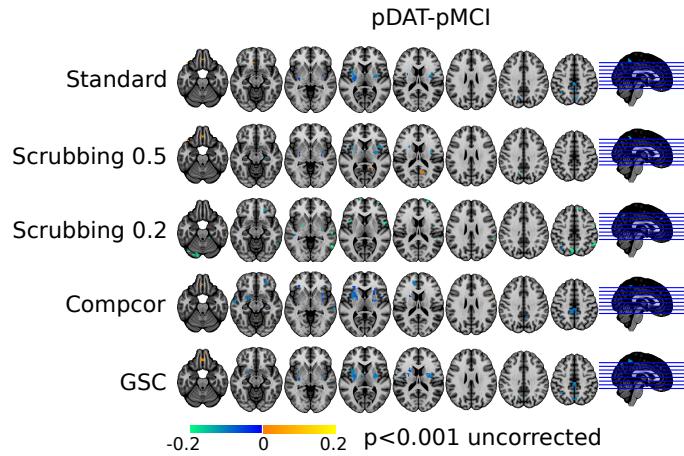


Figure 2.7: Impact of preprocessing on connectivity differences between the pDAT and pMCI groups. Connectivity differences between DMN (t -test $p<0.001$ uncorrected) computed with various preprocessing strategies (standard, scrubbing (0.5 and 0.2), CompCor and GSC). The maps are represented on top of the ICBM 152 anatomical atlas.

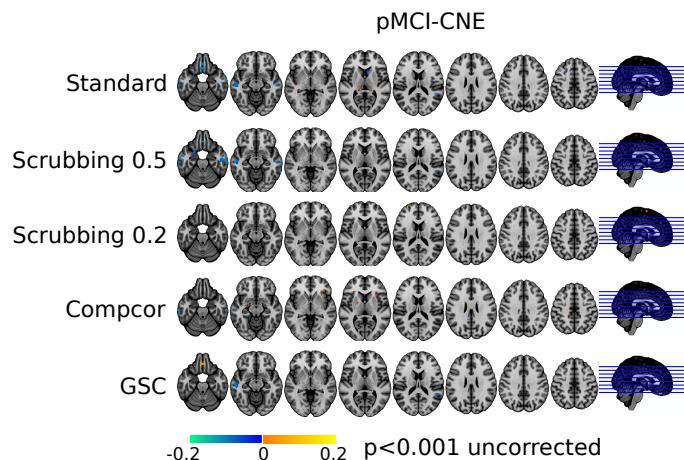


Figure 2.8: Impact of preprocessing on connectivity differences between the pMCI and CNE groups. Connectivity differences between DMN (t -test $p<0.001$ uncorrected) computed with various preprocessing strategies (standard, scrubbing (0.5 and 0.2), CompCor and GSC). The maps are represented on top of the ICBM 152 anatomical atlas.

2.3.4 Impact of scrubbing on our discriminative power between populations

2.3.4.0.3 point to point connections Despite a moderate loss of subjects using a scrubbing at 0.5 (due to insufficient remaining data < 50 frames), the gains in data quality translated into an increased detection rate for group differences in almost all tested connections. As shown in Figure 2.9 the $FD \geq 0.5$ scrubbing procedure mitigated motion artefacts and improved or maintain the statistical power of cross-sectional comparison of elderly clinical cohort. For $FD \geq 0.2$ some improvement can be observed but the sample size remain too small to be significant in most cases. The contrast with the most improvement due to scrubbing is the pDAT-CNE which is the contrast on which the literature point to point connections was selected for. For the pDAT-CNE contrast the most improved and consistent connections are connections with the IPL and the right SFG as well as the IPL and the dMPFC3. Connection with the PCC are also markedly improved by scrubbing namely the PCC and PCUN with MTL one of the first connection reported to be affected in the earliest stages of Alzheimer disease. Note that not all connection who had a good consistency with standard preprocessing improved using scrubbing. The statistical power for the pMCI-CNE contrast showed small improvements in PCC-PCUN, aMPFC-PCUN, IPL-dMPFC3 for $FD \geq 0.5$ and improved consistency for SFGr-FUS with a scrubbing at $FD \geq 0.2$. Finally the pDAT-pMCI contrast showed a general decrease in statistical power when scrubbing was applied, the only exception being aMPFC-PCUN that show improvement when a scrubbing at $FD \geq 0.2$ was applied.

2.3.4.0.4 Detection power for every scrubbing strategy The simulation applied using CompCor and GSC preprocessing strategies revealed less consistency with the previously mentioned pairs of connections, especially for the pDAT-CNE contrast, that was supposed to be the optimized contrast with that selection of connection pairs. For the pMCI-CNE contrast, GSC was consistently outperformed by CompCor and Standard preprocessing in the most dominant connection, namely dMPFC-dMPFC2, PCC-PCUNm and aMPFC-MTL. On the other hand, in the last contrast pDAT-pMCI, the CompCor preprocessing was the preprocessing strategy with the most improved statis-

tical power for aMPFC-PCUN, PCC-MTL and IPL-MTL. Interestingly, the scrubbing procedure seemed to improve the detection power more consistently and over a larger number of pair of connections than the CompCor and GSC methods (see Figure 2.9 and 2.10).

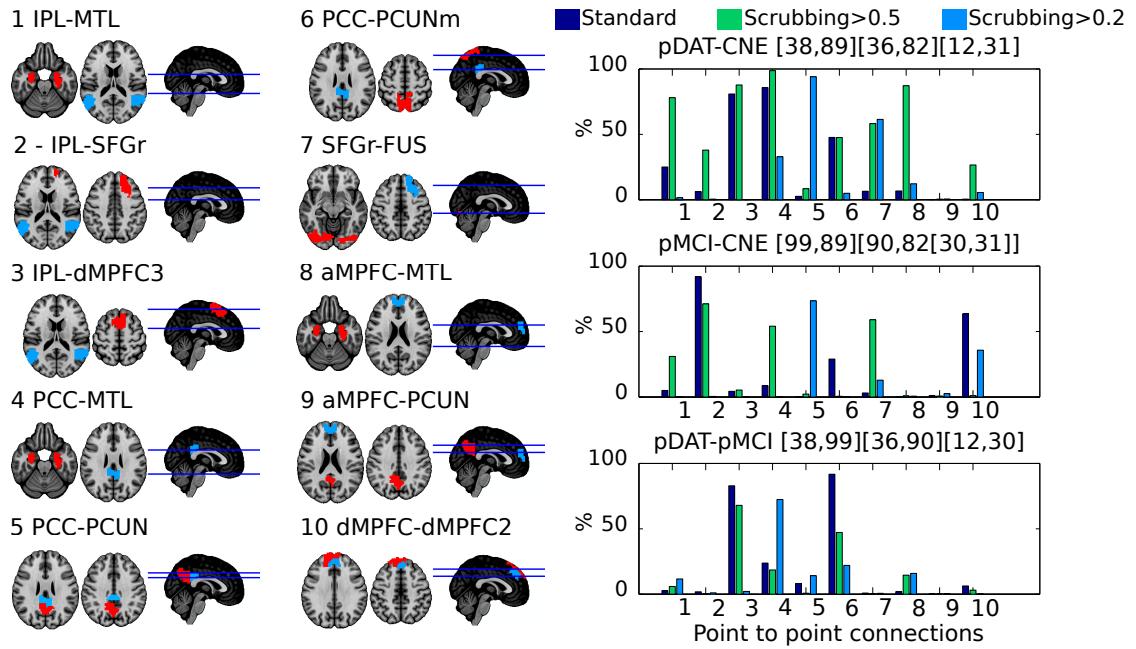


Figure 2.9: On the Left: Seeds from 10 point-to-point connections. PCC (posterior cingulate cortex), PCUN (precuneus), dMPFC (dorsomedial prefrontal cortex), dMPFC2 (dorsomedial prefrontal cortex2), IPL (inferior parietal lobule), SFGr (right superior frontal gyrus), aMPFC (anterior medial prefrontal cortex), PCUNm (precuneus motor), dMPFC3 (dorsomedial prefrontal cortex3), MTL (Mesial temporal lobe). On the right: Detection power of group differences for 3 preprocessing strategy (standard, scrubbing $FD > 0.5$ and scrubbing $FD > 0.2$). The detection power is computed using a t -test of each connection ($p < 0.05$) replicated $B = 10^4$ times using random subsamples of 70% of each group. Explanatory variables included age and gender and a multi-site bias correction are applied.

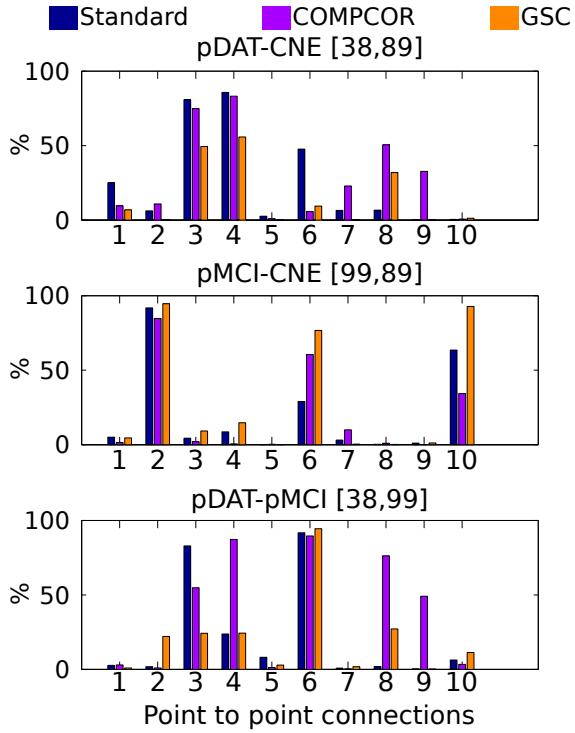


Figure 2.10: Detection power of group differences for 3 preprocessing strategy (standard, CompCor and GSC). The detection power is computed using a t -test of each connection ($p < 0.05$) replicated 10,000 times using random subsamples of 70% of each group. Explanatory variables included age and gender and a multi-site bias correction are applied.

2.4 Conclusion

We observed that motion introduced a systematic bias in the measures of resting-state connectivity in elderly population. Populations suffering did not exhibit more severe motion level than cognitively normal elderly subjects. The scrubbing procedure mitigated motion artefacts and improved the statistical power of the detection of differences in average functional connectivity between clinical populations.

CHAPTER 3

FEASIBILITY OF MULTI-CENTRIC FMRI CONNECTIVITY STUDIES OF ALZHEIMER'S DISEASE

My second on-going project investigates the multi-site bias in connectivity measures, and the methods that can be used to reduce this bias. Part of this work was published in the proceedings of the Alzheimer's Association International Conference (AAIC) Dansereau et al. (2013), and a full-length manuscript is currently in preparation.

3.1 Introduction

Resting-state (RS) connectivity in fMRI is a promising biomarker for a variety of neurological diseases. Typically, in a clinical trial, a large cohort is recruited and evaluated at multiple sites spread over countries or even continents. The main potential issue with that approach is the lack of consistency in the multi-site RS connectivity acquisitions that may obscure clinically relevant information. Therefore the aims of the study were to: (1) characterize the amplitude of the site bias, i.e. the systematic differences in rs-fMRI connectivity across different acquisition sites; (2) Quantify the impact of the between-site variance on the power of statistical tests in resting-state fMRI.

3.2 Method

3.2.1 Preprocessing

The datasets were analysed using the NeuroImaging Analysis Kit (NIAK¹) version 0.12.14, under CentOS version 6.3 with Octave² version 3.8.1 and the Minc toolkit³ ver-

¹<http://www.nitrc.org/projects/niak/>

²<http://gnu.octave.org>

³<http://www.bic.mni.mcgill.ca/ServicesSoftware/ServicesSoftwareMincToolKit>

sion 0.3.18. Analyses were executed in parallel on the "Mammouth" supercomputer⁴, using the pipeline system for Octave and Matlab (Bellec et al. 2010a), version 1.0.2. Brain map visualizations were created using MRICron software Rorden et al. (2007). Each fMRI dataset was corrected of inter-slice difference in acquisition time and the parameters of a rigid-body motion was estimated for each time frame. Rigid-body motion was estimated within as well as between runs, using the median volume of the first run as a target. The median volume of one selected fMRI run for each subject was coregistered with a T1 individual scan using Minctracc (Collins et al. 1998), which was itself non-linearly transformed to the Montreal Neurological Institute (MNI) template (Fonov et al. 2011) using the CIVET pipeline (Zijdenbos et al. 2002). The MNI symmetric template was generated from the ICBM152 sample of 152 young adults, after 40 iterations of non-linear coregistration. The rigid-body transform, fMRI-to-T1 transform and T1-to-stereotaxic transform were all combined, and the functional volumes were resampled in the MNI space at a 3 mm isotropic resolution. The a censoring method described in (Power et al. 2012) called "scrubbing" was used to remove the volumes with excessive motion using a cut-off value of $FD \geq 0.5$. A minimum number of 50 unscrubbed volumes per run, corresponding to ~ 125 s of acquisition for a TR of 2.5 seconds, was then required for further analysis. The following nuisance parameters were regressed out from the time series at each voxel: slow time drifts (basis of discrete cosines with a 0.01 Hz high-pass cut-off), average signals in conservative masks of the white matter and the lateral ventricles as well as the first principal components (95% energy) of the six rigid-body motion parameters and their squares (Lund et al. 2006),(Giove et al. 2009). The fMRI volumes were finally spatially smoothed with a 6 mm isotropic Gaussian blurring kernel.

3.2.2 Feature selection

Regions are routinely defined using an anatomical parcellation (He et al. 2009), such as the AAL template (Tzourio-Mazoyer et al. 2002). Anatomical parcels may however

⁴<http://www.calculquebec.ca/index.php/en/resources/compute-servers/mammouth-parallel-ii>

not well match the organization of resting-state networks. The BASC method was used to generate data-driven functional decomposition into resting-state networks based on the coherence of BOLD activity at the individual or group level (Bellec 2006; 2013, Bellec et al. 2010b). When a low number of networks (or scale) is used, the brain got decomposed into distributed large-scale networks, such as the DMN. At high scales (large number of networks), the BASC identified subnetworks and functional regions (Kelly et al. 2012). We generated a BASC parcellation in 100 clusters on the Cambridge sample, including ~ 200 young adults from the 1000 functional connectome database (Biswal et al. 2010) and used it to generate the rs-fMRI outcome measures.

3.2.2.0.5 The default-mode network Since the seminal work of Greicius et al. (2004), many rs-fMRI studies in AD focused on the default-mode network (DMN), a group of regions consistently more active at rest than during a broad range of different tasks (Gusnard and Raichle 2001). The DMN was notably reported to largely overlap with the regions that show high amyloid-beta deposition in patients with DAT (Buckner et al. 2009). It includes the posterior cingulate cortex (PCC) / Precuneus (PCUN) area, the inferior parietal lobule (IPL), the anterior cingulate cortex / medial prefrontal cortex (MPFC) (Greicius et al. 2003). Other structures such as the medial temporal cortex or the superior frontal gyrus are also generally regarded as part of different subnetworks of the DMN (Andrews-Hanna et al. 2010, Margulies et al. 2009).

3.2.2.0.6 Literature review: Alzheimer’s disease and resting-state fMRI We performed a literature review to select candidate connections that have been shown to be prominently impacted in Alzheimer’s disease. There is no single authoritative reference on the effect of a DAT on rs-fMRI connectivity, and the field has been dominated thus far by studies with small samples ($n \sim 20$) and limited statistical power, see Sheline and Raichle (2013) for a recent review. Because the DMN has been most extensively studied, we decided to focus on this network and to run a meta-analysis on six papers that (1) used some analogue of seed-based connectivity maps in resting-state fMRI using one or multiple seeds in the DMN (2) investigated abnormalities in resting-state functional

connectivity in patients suffering of a dementia of the Alzheimer's type and (3) provided tables of coordinates in stereotaxic space for the results.

3.2.2.0.7 review: Alzheimer's disease and resting-state fMRI

- Zhang et al. (2009) used functional connectivity maps with a seed in the posterior cingulate cortex (PCC) to explore the differences between a group of elderly cognitively normal subjects (CNE, n=16) and patients with a mild dementia of the Alzheimer's type (DAT, n=18).
- Zhang et al. (2010) generalized the Zhang et al. (2009) study with CNE (n=16) and a larger group of patients with DAT (n=46). Patients were separated in three groups (mild, moderate, severe DAT), and each group of patients was contrasted against the CNE.
- Wang et al. (2006) used functional connectivity maps with a seed in the hippocampi to explore the differences between a group of CNE (n=13) and patients with a mild DAT (n=13). All results included in the meta-analysis are from Table 2, seeded in the right hippocampus. Seeds were manually delineated on an individual basis.
- Wang et al. (2007) used functional connectivity maps with a seed in the posterior cingulate cortex (PCC) as well as full brain point-to-point correlations (based on an AAL parcellation) to explore the differences between a group of elderly cognitively normal subjects (CNE, n=14) and patients with a very mild to mild dementia of the Alzheimer's type (DAT, n=14). Only the results based on the PCC seed were included in the meta-analysis.
- Goveas et al. (2011) used functional connectivity maps with a seed in the hippocampi to explore the differences between a group of elderly cognitively normal subjects (CNE, n=18) and patients with a mild dementia of the Alzheimer's type (DAT, n=14) before and after donepezil treatment. Seeds were manually delineated on an individual basis, before and after treatment.

- Damoiseaux et al. (2012) used dual-regression independent component analysis to explore longitudinal differences between a group of CNE (n=18) and patients with DAT (n=21). All results included in the meta-analysis are from Table 3 (differences at baseline) and Table 4 (interaction with time). The authors used three components representing the Anterior DMN, Ventral DMN and Posterior DMN.

To assess the degree of consistency of the findings across studies, we counted the number of coordinates located in each one of the BASC regions. The resulting map is presented in Figure 3.2. As can be seen, there is a lot of variability across studies, with only a limited number of regions reaching a score above 3 (i.e. reported in at least 3 of the contrasts in the six studies). Note that we did not select connections with the hippocampus, although this region was frequently reported. The rationale was that the drug effects on this area are expected to be minimal in patients with a moderate DAT, because the very severe atrophy of the structure cannot be recovered. In the regions showing the most consistency (score of 3 or more), there were many regions located in the DMN, such as the PCC, the PCUN, the IPL (a bilateral node), the right superior frontal gyrus (SFGr), as well as two dorsal MPFC cortex (dMPFC and dMPFC2) and an anterior MPFC parcel (aMPFC). Three parcels were found in the visual network: the lingual gyrus (LING), the fusiform gyrus (FUS) and a dorso-medial occipital (DMO) parcel. Two parcels were found in the dorsal attentional network: the intra-parietal sulcus (IPS) and the motor part of the precuneus (PCUNm, see Margulies et al. (2009)). One parcel was found in the premotor cortex (PMC), associated with the sensorimotor network, one parcel in the left dorsolateral prefrontal cortex (rDLPFC), associated with the fronto-parietal task-control network, as well as a parcel in the dMPFC (dMPFC3) associated with the cingulo-opercular cortex. Finally, a parcel included the temporal poles (TPo) bilaterally. Note that the nomenclature for distributed networks was based on (Power et al. 2011).

The TRT reliability study was based on the publicly available NYU-TRT database. The database included 25 young healthy adults, and each subject had three rs-fMRI run: two in a single session (separated by 45 mns) and another run 5 – 16 months latter. Several outcome measures were generated in key regions impacted by AD. Using the

three runs, one intra-class correlation (ICC) was generated intra-session, and two ICCs were generated inter-session for each outcome measures. The outcome measures were ranked based on average of intra- and inter-session ICCs.

3.2.3 Simulations

3.2.3.0.8 Dataset In order to simulate various scenarios within the context of a multi-site setting, a cohort of subjects acquired at a single-site was selected to act as our reference dataset and for the multi-site configuration a cohort from a collection of 7 small sites, roughly totalling the same sample size as the reference dataset, was used. The cohort used for the study contains 385 participants from the 1000 Functional Connectomes Project (Biswal et al. 2010) (150 males, age range = 18-46 yrs) composed of 1 large site (Cambridge n=~200) and 7 small sites (n=~20/site for a total ~200). The fMRI datasets were preprocessed with the Neuro-Imaging Analysis Kit (NIAK) as described earlier in Section 3.2.1.

3.2.3.0.9 Connectome Using a brain partition of R networks obtain from BASC procedure described in Bellec et al. (2010b), and taking each pair of distinct networks i and j , the between-network connectivity $y_{i,j}$ is measured by the Fisher transform of the Pearson's correlation between the average time series of the network. The within-network connectivity $y_{i,i}$ is the Fisher transform of the average correlation between time series of every pair of distinct voxels inside network i . The connectome $\mathbf{Y} = (y_{i,j})_{i,j=1}^R$ is thus a $R \times R$ matrix. Each column j (or row, as the matrix is symmetric) codes for the connectivity between network j and all other brain networks (full brain functional connectivity map). For a scale with R parcels, there are exactly $L = R(R + 1)/2$ distinct elements in an individual connectome \mathbf{Y} .

3.2.3.0.10 Effect size (cohen's d) For each site and each sample, half of the subjects were randomly assigned to a "treatment" group and a Monte-Carlo simulation was used to estimate the detection power in the single-site and in multi-site setting.

The normalized Cohen's d was used to estimate the effect size and it is defined as the difference between two means \bar{x}_1, \bar{x}_2 divided by a standard deviation from the data s .

$$\begin{aligned} s &= \sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}} \\ d &= \frac{\bar{x}_1 - \bar{x}_2}{s} \end{aligned} \quad (3.1)$$

n_1 and n_2 are the respective number of subject in each group. In order to introduce the same effect-size across the single-site and multi-site dataset we are taking the standard deviation from the single-site cohort as the reference. The connection $y_{i,j}$ of the randomly affected subjects ("treatment" group) are therefore calculated $y_{i,j} = y_{i,j} + d \times s_{i,j}$. The significance of the difference between the control and 'treatment' group was assessed by a t -test in a linear model, including a covariate to model the motion. The study was repeated for various effect sizes (0 to 0.8 with a step of 0.01) with a p -value threshold of 0.05 on the t -test.

3.2.3.0.11 multi-site correction approaches For the multi-site three flavours were computed: multi-site no correction, multi-site with dummy variables and multi-site with METAL correction. Depending on the multi-site configuration and distribution of the subject we proposed two corrective approaches that can be applied as shown in the simulations of Figure 3.8. The first one is the introduction of dummy variables (binary vectors $1 \times N$) who code for each site in the GLM model 3.3.

The variables are corrected to have a zero mean across subjects, and an intercept (i.e. a column filled with 1) is added to \mathbf{X} . The GLM relies on the following stochastic model:

$$\mathbf{Y} = \mathbf{X}\beta + \mathbf{E} \quad (3.2)$$

where β is an unknown $C \times L$ matrix of linear regression coefficients and \mathbf{E} is a $N \times L$ random (noise) multivariate Gaussian variable. To apply the correction $v - 1$ dummy-variables are added to the model 3.3 with v being the total number of sites used in the study.

$$y_{i,j} = \beta x + \beta_{i=1,\dots,v-1} x_{i=1,\dots,v-1} + e \quad (3.3)$$

The second approach is to compute the GLM independently on each site and then combine the statistical results from each site in a global score. This model averaging technique called METAL from Willer et al. (2010) model site specific bias by running a GLM analysis on each site resulting in v beta vectors that are weighted proportionally to the standard error of each site and finally averaged as shown in equation 3.9. This is the most flexible way to account for multi-site effect while keeping the analysis simple and robust to unbalanced sites.

$$\beta_v, \text{effect size estimate for site } v. \quad (3.4)$$

$$se_v, \text{standard error for site } v. \quad (3.5)$$

$$w_v = \frac{1}{se_v^2}, \text{weight estimate for site } v. \quad (3.6)$$

$$SE = \sqrt{\frac{1}{\sum_v w_v}}, \text{global standard error.} \quad (3.7)$$

$$\beta = \frac{\sum_v \beta_v w_v}{\sum_v w_v}, \text{global } \beta. \quad (3.8)$$

$$Z = \frac{\beta}{SE}, \text{global Z score.} \quad (3.9)$$

$$p = 2(1 - \phi(|Z|)), \text{p-value.} \quad (3.10)$$

3.3 Results

The Figure 3.1 presents the results of the ICC analysis for the point-to-point correlations, only the connections with an average ICC above 0.5 are represented. The results were consistent with (Shehzad et al. 2009), with a mean ICC over all connections of ~0.3 and 23 connections scoring a moderate-to-good level of ICC (> 0.5).

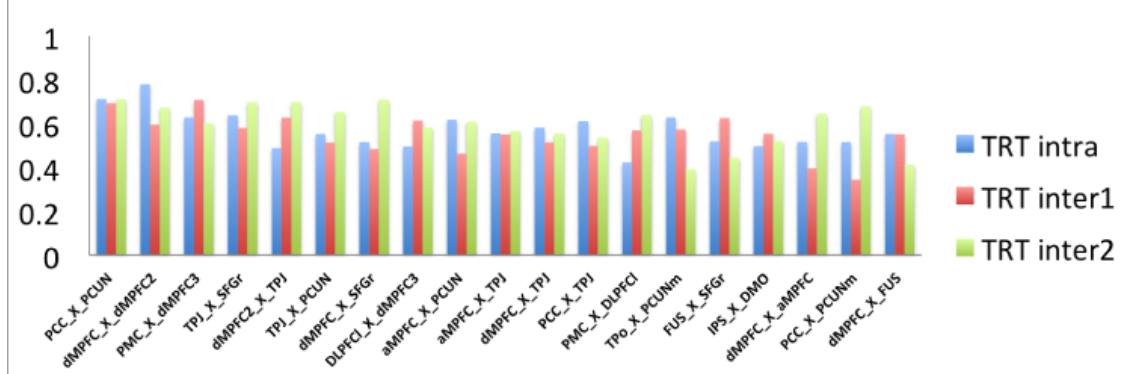


Figure 3.1: ICC scores for pairs of connections passing the > 0.5 threshold on the NYU-TRT dataset.

Average connectivity maps associated with the DMN nodes as well as the nodes outside the DMN that passed the TRT selection are presented in Figures 3.3 and 3.4. The involvement of the sensorimotor, visual and attentional networks mainly came from contrasts reporting a decrease of negative correlations in patients, that was interpreted as a compensation mechanism by some authors. These connections are potentially very valuable to monitor the effect of a drug. Considering that we pooled studies of the DMN, we decided to select as candidates all connections of parcels within the DMN, as well as connections between a parcel inside the DMN and a parcel outside the DMN. The final selection of target measures was based on test-retest reliability. All the parcels and associated labels and networks are listed in Table 3.I.

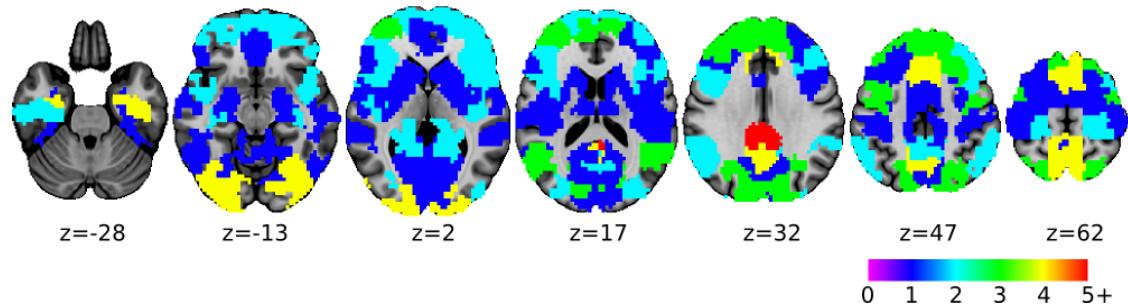


Figure 3.2: Frequency of reported regions showing functional differences based on a literature review of 6 papers.

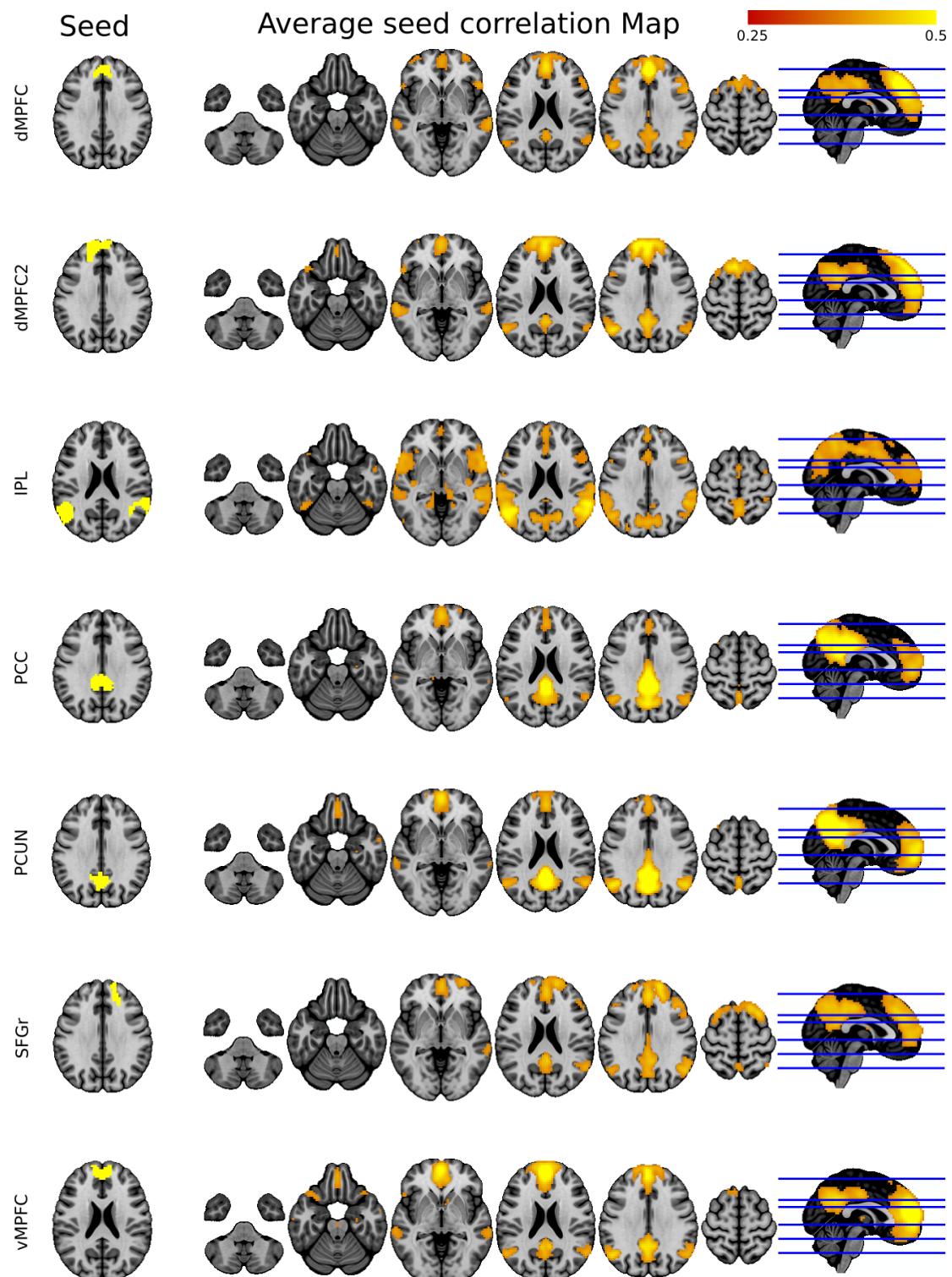


Figure 3.3: Selected nodes inside the DMN who passed the TRT selection.

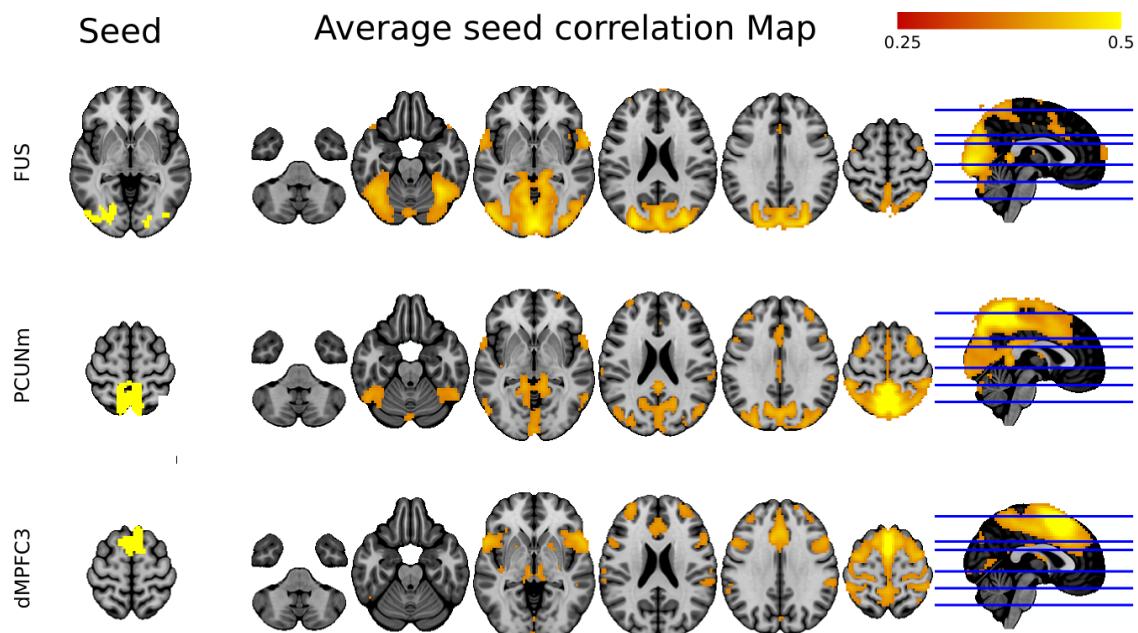


Figure 3.4: Selected nodes outside the DMN who passed the TRT selection.

Network	Label	Name	Cambridge100
Default-mode network	PCC	posterior cingulate cortex	1
	dMPFC	dorsomedial prefrontal cortex	12
	dMPFC2	dorsomedial prefrontal cortex	46
	aMPFC	anterior medial prefrontal cortex	42
	IPL	inferior parietal lobule	49
	PCUN	precuneus	53
Visual network	MTL	medial temporal lobe	39
	SFGr	right superior frontal gyrus	76
Visual network	FUS	fusiformgyrus	71
Dorsal attentional	PCUMm	precuneus (motor)	94
Cingulo-opercular network	dMPFC3	dorsomedial prefrontal cortex	90

Table 3.I: Regions selected in the literature review, the region number correspond to the number in the Cambridge 100 partition.

For point-to-point correlations within the DMN, we selected the connections with highest ICC for each node (all average $ICC > 0.5$):

- PCC x PCUN
- dMPFC x dMPFC2
- IPL x SFGr
- aMPFC x PCUN

For each point-to-point correlation between the DMN and another network, we selected the connections with highest ICC and $ICC > 0.5$:

- FUS x SFGr
- PCC x PCUNm
- IPL x dMPFC3

The first assessment perform on the dataset was to verify the distribution of the variance in functional connectivity among each site and across sites in order to see if they are of the same order of magnitude or not. This analysis of Figure 3.5 shows the distribution of the standard deviation of connectivity across subjects (the distribution is over the full brain connectome, with several 1000s connections) at the 8 sites against the inter-sites standard deviation of connectomes (average at each site). as we can see the inter-site (between site) variability is smaller than the intra-site (between subjects) variability.

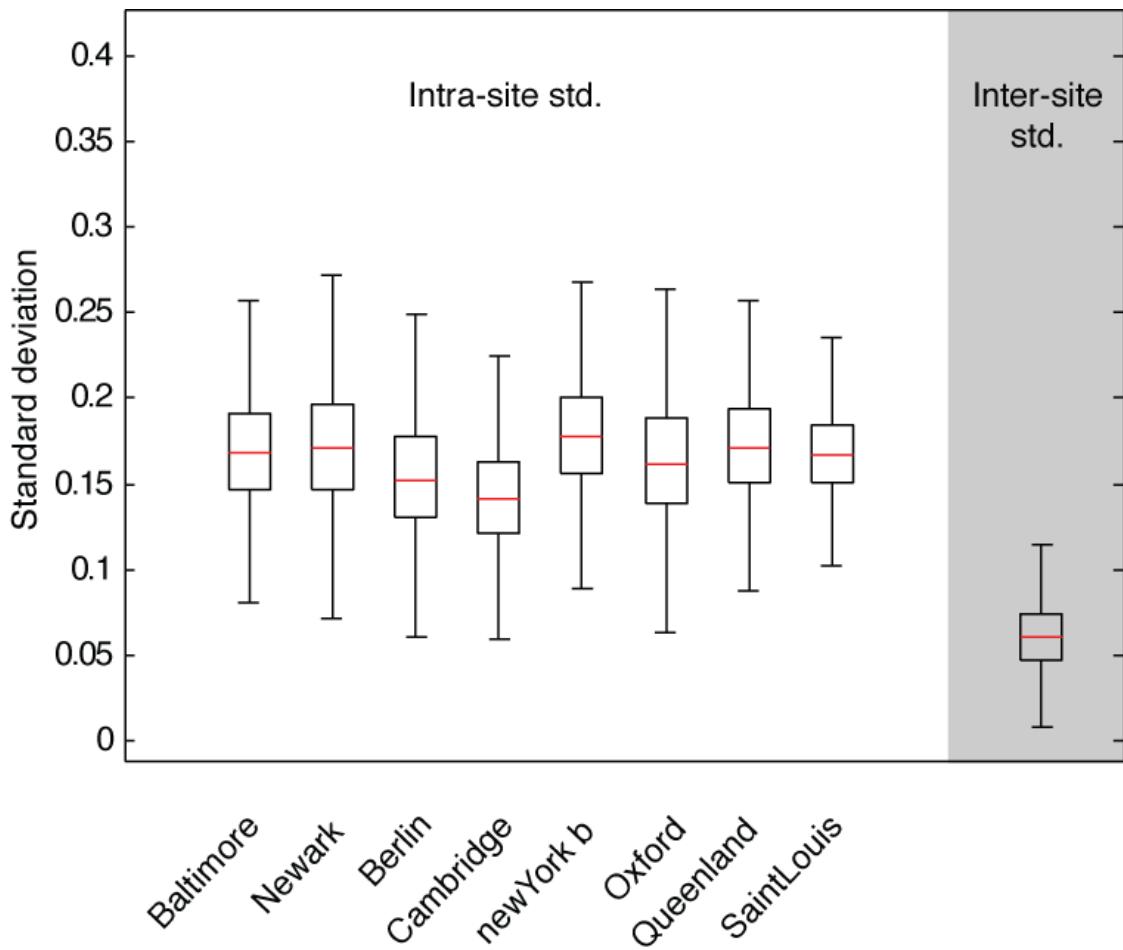


Figure 3.5: Distribution of intra-site (between-subject) standard deviation vs. inter-site (between-site) standard deviation, based on the standard deviation of the connectivity matrices from 8 sites from the 1000 functional connectome dataset.

In order to verify how spatial structure vary across sites the average standard deviation and the average connectivity map of the DMN was extracted for each site and displayed in Figure 3.6. In order to ease the reading we selected only 4 representative sites, although we reached the same conclusion on all the sites. As we can see in the intersection between two sites the difference in standard deviation between-sites was illustrated (red set of brain cuts). First the mean DMN at each site is consistent with the expected spatial ditribution reported in other studies. As we can see the amplitude of inter-site bias is about 3-fold smaller than the within-site standard deviation (red ~0.06 vs. orange ~0.18). The most salient changes between-sites are located in the mesio-frontal region associated with the anterior part of the DMN. This last finding may be associated with motion artefact as previously reported in Dansereau et al. (2014b).

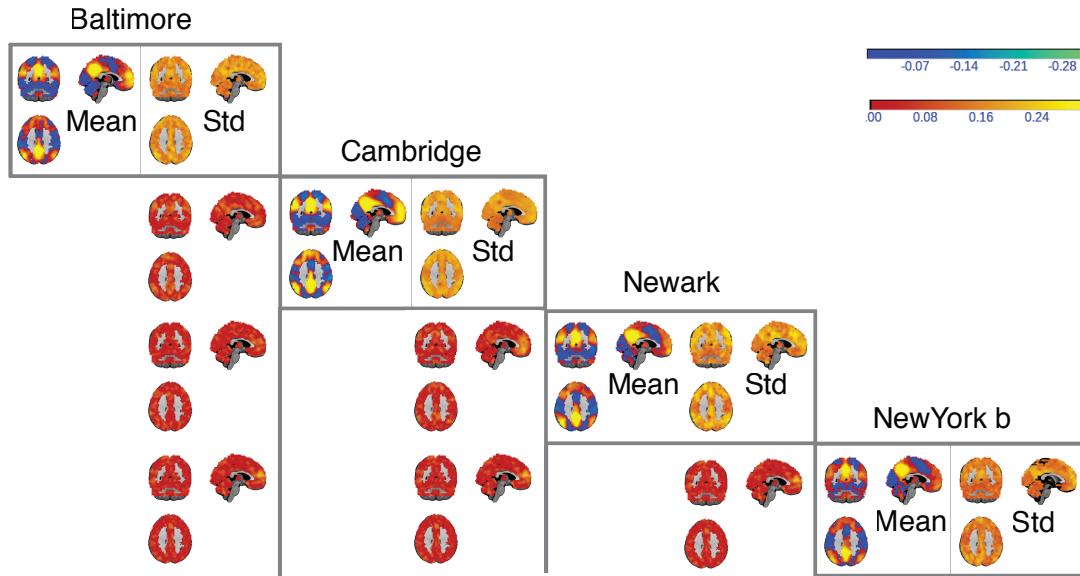


Figure 3.6: Functional connectivity maps of the default-mode network at multiple sites. The average connectivity map for 4 sites (Baltimore, Cambridge, Newark and New-York b at 3T) are presented on the diagonal (left). The standard deviation across subjects and within site is presented next to it (diagonal squares, right part). Each off-diagonal block represent the absolute difference between the average functional connectivity maps between two sites (called the inter-site bias).

We also showed using Monte-Carlo simulations that the power of detecting an effect is marginally affected by the site acquisition configuration (single site or multi-site, see Figure 3.2 for an illustration of a power analysis on three different seeds) where the sites are balanced in term of the amount of subject with and without the effect.

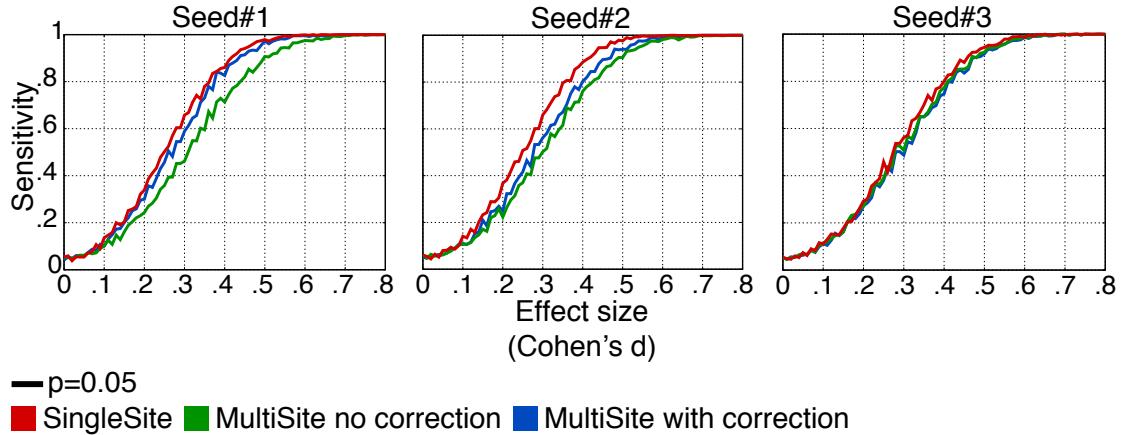


Figure 3.7: Power analysis for a resting-state fMRI study. A Monte-Carlo simulation was implemented to evaluate the power of a resting-state multi-site study, based on real values of three connections in the default-mode network PCC/MFC, rHIP/vMFC and lIIPC-rDLPFC. For each site and each sample, half of the subjects were randomly assigned to a 'treatment' group. For the subjects in this group, a value was added to achieve a given relative effect size (Cohen's d , i.e. the mean of the two groups divided by the standard deviation of all sites). The significance of the difference between the control and 'treatment' group was assessed by a t -test in a linear model, including covariates to model site-specific bias. The study was repeated for various effect sizes (0 to 0.8 with a step of 0.01) at a threshold of 0.05 on the p -value in the t -test. For a p of 0.05, a statistical power of 0.95 can be achieved for as low as a 0.47 effect size. The simulation was based on a scenario with 8 sites and 385 subjects, and no homogenization of acquisition protocol whatsoever. The multi-site (with and without correction) is based on 187 subjects from 7 sites and the SingleSite is based on one site of 198 subjects.

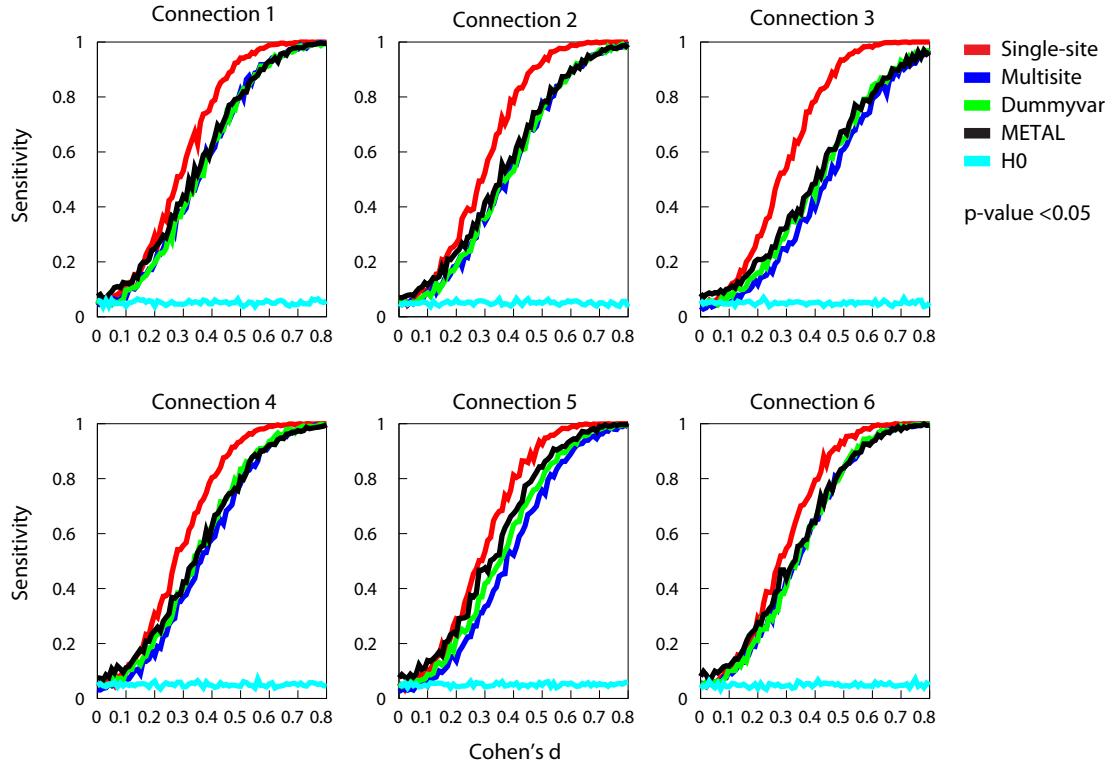


Figure 3.8: Power analysis for a resting-state fMRI study. A Monte-Carlo simulation was implemented to evaluate the power of a resting-state multi-site study. For each site and each sample, half of the subjects were randomly assigned to a 'treatment' group. For the subjects in this group, a value was added to achieve a given relative effect size (Cohen's d , i.e. the mean of the two groups divided by the standard deviation of all sites). The significance of the difference between the control and 'treatment' group was assessed by a t -test in a linear model. To correct for site-specific bias two correction are presented the dummy variables and METAL. The study was repeated for various effect sizes (0 to 0.8 with a step of 0.01) at a threshold of 0.05 on the p-value in the t -test. The simulation was based on a scenario with 8 sites and 385 subjects, and no homogenization of acquisition protocol whatsoever. The multi-site (with and without correction) is based on 187 subjects from 7 sites and the Single-site is based on one site of 198 subjects.

3.4 Conclusion

We confirmed that multi-site acquisition introduce some variability in the dataset although a single-site study with 200 subjects had only a marginally superior statistical power than an analysis pooling 7 sites for an equivalent number of subjects, see Figure 3.7. In both cases, a high sensitivity (> 0.95) could be achieved for the effect size observed by Goveas et al. (2011) which reported an effect equivalent to 1. We can therefore conclude that it is feasible to acquire rs-fMRI data in across multiple sites and correct for this topology.

CHAPTER 4

FUTURE DEVELOPMENT

4.1 The connectome feature space

Once the data is preprocessed properly, we need to derive some metric of brain connectivity. One of the most common metric is the Pearson correlation, and is usually represented as a $n \times n$ connectivity matrix of every voxel in the gray matter. As there are in the order of 10^4 voxels, there is an overwhelming number of 5×10^7 possible connections to examine as a potential diagnostic tool. We therefore have a very large number of features, but only few of them are potentially relevant for diagnostic purposes. Unfortunately, only a quite limited number of examples are available in a typical neuroimaging study to learn which connections are the most informative ($< 10^3$). Even state-of-the-art classification algorithms (e.g. SVM (Cortes and Vapnik 1995)) cannot overcome the presence of large number of weakly relevant and redundant features. This is usually attributed to "the curse of dimensionality" (Bellman and Bellman 1961), or to the fact that irrelevant features decrease the ability of the learner to discriminate between classes. Moreover, many machine-learning algorithms become computationally intractable when the dimension is high. On the other hand once a reduce set of features has been chosen even the most basic classifiers can achieve high performance classification.

In order to reduce that feature space and to have a clinically meaningful representation of functional structures, a number of solutions have been proposed that take advantage of the underlying structure of the brain (van den Heuvel et al. 2009). Regions are routinely defined using an anatomical parcellation (He et al. 2009), such as the AAL template (Tzourio-Mazoyer et al. 2002). Anatomical parcels may however not well match the organization of resting-state networks. We use a framework to generate data-driven functional decomposition into resting-state networks based on the coherence of fMRI time series at the individual or group level (Bellec 2006, Bellec et al. 2010b). When a low number of networks (or scale) are used, this technique, called bootstrap analysis

of stable clusters (BASC), generates decompositions of the brain into distributed large-scale networks, such as the default mode network (DMN). At high scales, it identifies sub-networks and functional regions (Kelly et al. 2012). We can therefore use those parcellation units to reduce the feature space. Correlation is a good approach to asses connectivity but can be highly variable in noisy dataset. This motivated me to investigate another metric called "stability" based on evidence accumulation of clustering on bootstrap samples, that could potentially be more consistent across subjects and scanning session resulting in improved prediction power and generalizability. The stability metric was design in an attempt to reduce the variance in the acquisition and extract the most consistent functional structures in the underling data by bootstrapping temporal blocks to identify the regions that are most often clustered together even though the temporal block are re-sampled. In an attempt to extend my current work I will use the simulations that I have design to evaluate the multi-site effect (1000 functional connectome dataset composed of one large site and 7 smaller sites) in order to evaluate if the detection power is improved or not using a stability metric instead of a correlation metric and if it is robust to various multi-site scenarios.

4.2 Feature selection and classification

A crucial point of our study is to identify the most discriminant features to use in the prediction model and account for various confounding factors (e.g. age, gender, education, multi-site effect and motion). This is future work that will be conducted in the next year and a half. I will particularly focus on the stability of the features selected, in order to obtain robust and consistent markers of the disease. In order to estimate stability of the selection I plan to use a Monte-Carlo estimation of the selection using bootstrap resampling (Bellec et al. 2010b, Efron and Tibshirani 1994) on the training dataset inside a 10-fold cross-validation. The large feature space that will be fed to this procedure will be a correlation matrix of network at multiple scales. As an example, let's take 3 parcellations (e.g. scales 10, 50, 100) of the functional brains obtained from an independent dataset of normal subjects using the method describe in the previous

section (Large feature space). Using the time-series of every parcellation I will obtain a correlation matrix A of size $R \times R$ and $R = 10 + 50 + 100 = 160$ resulting in a vectorized form of this matrix of size $L = R(R + 1)/2 = 12880$ unique features. The idea behind this strategy is to capture interaction of small networks with larger networks instead of just looking at the interaction of large network with large network or small network with small network which could miss some important interaction. An example of previously mentioned interaction of larger network with smaller ones is the known decrease in connectivity between the hippocampal structure (small regions) with the default mode (a large and extended network). This procedure could potentially capture those interactions that would in term maximize the prediction accuracy as well as controlling for stability by only retaining the most stable features identified by the feature selection therefore enforcing generalizability. Venkataraman et al. (2010) used stability measures with Gini importance metric and found a good performance distinguishing between patients with schizophrenia and normal control.

More work also needs to be done in finding the appropriate classification method and investigating several standard alternatives in conjunction with novel feature selection methods will also be part of my project in the next year. I will notably consider the Linear Discriminant Analysis (LDA) because it is straightforward to include covariates in the model in order to account for confounding effects. I will also evaluate SVM since it is a very popular method, and I will more particularly investigate the margin optimisation criteria for feature selection (Gilad-Bachrach et al. 2004, Kononenko et al. 1997). One last approach will be to use ensemble techniques like AdaBoost (Freund and Schapire 1997) to improve generalization performance. AdaBoost has been proved to be remarkably resistant to overfitting (Schapire et al. 1998). An interesting point is that Schapire et al. (1998) has shown using a slightly different definition of the margin that AdaBoost also "boosts" the margins, meaning that it finds the decision boundary that is further away from the instances of all classes. By margin Schapire refers to the difference between the total vote AdaBoost receives from correctly identifying classifiers and the maximum vote received by any incorrect class.

The evaluation of the classification pipeline will be done using two datasets namely:

1) the Cobre dataset composed of 74 control and 72 schizophrenic subjects and 2) the enhanced Nathan Kline Institute-Rockland Sample (NKI-RS) a large community sample representing > 500 subjects. The Cobre dataset will be used to study the ability of the method to identify discriminative features of schizophrenia (previous univariate analysis have shown large effect on the same dataset). We will also use the NKI-RS to predict the effect of age. Another interesting aspect of the NKI-RS is that every subject has been scanned with different fMRI protocol (scanner parameters) on the same scanner. We can use this information to have a pseudo multi-site data set and explore the generalization of the trained classifier on the same with a different scanning protocol.

4.3 Industry application and translational effort

I am involved in the industry (through my consulting work with the companies NeuroRX and Biospective), I'm advising on the fMRI analysis of multiple clinical trials and looking at the feasibility of using fMRI in multicentric pharmaceutical trials (some of these work have lead to publications). I'm also doing statistical analysis and proposing biomarkers tailored to the clinical questions of the pharmaceutical sponsors. I'm also starting to get involved with the biomarker unit of the Canadian Consortium on Neurodegeneration in Aging (CCNA) who wants to propose new biomarker for Alzheimer disease that would be used systematically in the clinical assessment of Alzheimer disease across Canada. These efforts and collaborations are perfectly in line with the objectives of my PhD. and are greatly contributing to translational findings and application of my research in pharmaceutical trials as well as addressing concrete question in the field of neuroimaging.

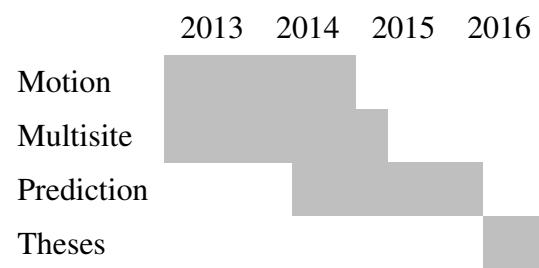
4.4 Timeline

I have identified three major points that I will address in my PhD:

- Investigate motion impact on functional connectivity I'm currently writing a manuscript that should be submitted at the end of this year (2014) on that specific topic. The

results have been presented (poster format) at the Organization for Human Brain Mapping conference 2013 (OHBM) and at the Alzheimer's Association International Conference 2014 (AAIC).

- Feasibility of multi-site connectivity analysis (inter-site normalization) Most of the analyses are completed and some of the results have been presented in two conferences (poster format), namely OHBM2013 and AAIC2013. I'm planning to submit the manuscript for this study mid 2015.
- Prediction
 - I'm currently experimenting with some standard machine learning tools and I will start by testing the pipeline on simple simulations and verify is the stability metric is more resistant to structured noise.
 - Then on a real dataset I will apply the pipeline on Cobre and NKI-RS.
 - Finally I will apply the pipeline on a dataset that I have compiled in the past year, this dataset is composed of 313 elderly adults with and without cognitive impairment of the Alzheimer type collected across 5 studies: ADNI2 study and 4 other studies based in Montreal, Canada, for a grand total of 126 CNE participants, 133 patients with MCI, and 54 patients with DAT. I will try to classify the various population in a cross sectional analysis. I am also planning to use a subset of the data namely the ADNI2 dataset who is a longitudinal study to assess the potential of the method to predict time of conversion from pMCI to pDAT. I plan to be done with those analyses in the winter of 2016 and publish the results in the following months.
- Finally I plan to submit my theses at the end of 2016 (see Timeline 4.4)



REFERENCES

- A. Anand, Y. Li, Y. Wang, J. Wu, S. Gao, L. Bukhari, V. P. Mathews, A. Kalnin, and M. J. Lowe. Activity and connectivity of brain mood regulating circuit in depression: a functional magnetic resonance study. *Biological psychiatry*, 57(10):1079–1088, May 2005. ISSN 0006-3223. doi: 10.1016/j.biopsych.2005.02.021. URL <http://dx.doi.org/10.1016/j.biopsych.2005.02.021>.
- Jessica R. Andrews-Hanna, Jay S. Reidler, Jorge Sepulcre, Renee Poulin, and Randy L. Buckner. Functional-anatomic fractionation of the brain’s default network. *Neuron*, 65(4):550–562, February 2010. ISSN 1097-4199. doi: 10.1016/j.neuron.2010.02.005. URL <http://dx.doi.org/10.1016/j.neuron.2010.02.005>.
- R. Baumgartner, C. Windischberger, and E. Moser. Quantification in Functional Magnetic Resonance Imaging: Fuzzy Clustering vs. Correlation Analysis. *Magnetic Resonance Imaging*, 16(2):115–125, February 1998. doi: 10.1016/S0730-725X(97)00277-4. URL [http://dx.doi.org/10.1016/S0730-725X\(97\)00277-4](http://dx.doi.org/10.1016/S0730-725X(97)00277-4).
- Yashar Behzadi, Khaled Restom, Joy Liau, and Thomas T. Liu. A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. *NeuroImage*, 37(1):90–101, August 2007. ISSN 1053-8119. doi: 10.1016/j.neuroimage.2007.04.042. URL <http://dx.doi.org/10.1016/j.neuroimage.2007.04.042>.
- P. Bellec, F. Carbonell, V. Perlberg, and A. C. Evans. *A neuroimaging analysis kit for Octave and Matlab*, 2010a. URL <http://code.google.com/p/niak/>.
- Pierre Bellec. *Longitudinal study of large-scale networks in the human brain using fMRI : Methods and application to motor skill learning*. PhD thesis, Université Paris XI, Orsay, Paris, February 2006. URL http://exocet.imec.jussieu.fr/~{}pbellec/data/these_bellec.pdf.

Pierre Bellec. Mining the hierarchy of resting-state brain networks: selection of representative clusters in a multiscale structure. In *Pattern Recognition in Neuroimaging (PRNI), 2013 International Workshop on*, pages 54–57. IEEE, 2013.

Pierre Bellec, Pedro Rosa-Neto, Oliver C. Lyttelton, Habib Benali, and Alan C. Evans. Multi-level bootstrap analysis of stable clusters in resting-state fMRI. *NeuroImage*, 51(3):1126–1139, July 2010b. ISSN 1095-9572. doi: 10.1016/j.neuroimage.2010.02.082. URL <http://dx.doi.org/10.1016/j.neuroimage.2010.02.082>.

Pierre Bellec, Sébastien Lavoie-Courchesne, Phil Dickinson, Jason P. Lerch, Alex P. Zijdenbos, and Alan C. Evans. The pipeline system for Octave and Matlab (PSOM): a lightweight scripting framework and execution engine for scientific workflows. *Frontiers in neuroinformatics*, 6, 2012. ISSN 1662-5196. doi: 10.3389/fninf.2012.00007. URL <http://dx.doi.org/10.3389/fninf.2012.00007>.

R. Bellman and R.E. Bellman. *Adaptive Control Processes: A Guided Tour*. 'Rand Corporation. Research studies. Princeton University Press, 1961. URL <http://books.google.ca/books?id=POAmAAAAMAAJ>.

B. Biswal, F. Z. Yetkin, V. M. Haughton, and J. S. Hyde. Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. *Magnetic resonance in medicine*, 34(4):537–541, October 1995. ISSN 0740-3194. URL <http://view.ncbi.nlm.nih.gov/pubmed/8524021>.

Bharat B. Biswal, Maarten Mennes, Xi-Nian N. Zuo, Suril Gohel, Clare Kelly, Steve M. Smith, Christian F. Beckmann, Jonathan S. Adelstein, Randy L. Buckner, Stan Colcombe, Anne-Marie M. Dogonowski, Monique Ernst, Damien Fair, Michelle Hampson, Matthew J. Hoptman, James S. Hyde, Vesa J. Kiviniemi, Rolf Kötter, Shi-Jiang J. Li, Ching-Po P. Lin, Mark J. Lowe, Clare Mackay, David J. Madden, Kristoffer H. Madsen, Daniel S. Margulies, Helen S. Mayberg, Katie McMahon, Christopher S. Monk, Stewart H. Mostofsky, Bonnie J. Nagel, James J. Pekar, Scott J. Peltier, Steven E. Petersen, Valentin Riedl, Serge A. Rombouts, Bart Rypma, Bradley L.

Schlaggar, Sein Schmidt, Rachael D. Seidler, Greg J. Siegle, Christian Sorg, Gao-Jun J. Teng, Juha Veijola, Arno Villringer, Martin Walter, Lihong Wang, Xu-Chu C. Weng, Susan Whitfield-Gabrieli, Peter Williamson, Christian Windischberger, Yu-Feng F. Zang, Hong-Ying Y. Zhang, F. Xavier Castellanos, and Michael P. Milham. Toward discovery science of human brain function. *Proceedings of the National Academy of Sciences of the United States of America*, 107(10):4734–4739, March 2010. ISSN 1091-6490. doi: 10.1073/pnas.0911855107. URL <http://dx.doi.org/10.1073/pnas.0911855107>.

Randy L. Buckner, Jessica R. Andrews-Hanna, and Daniel L. Schacter. The Brain’s Default Network. *Annals of the New York Academy of Sciences*, 1124(1):1–38, March 2008. ISSN 0077-8923. doi: 10.1196/annals.1440.011. URL <http://dx.doi.org/10.1196/annals.1440.011>.

Randy L. Buckner, Jorge Sepulcre, Tanveer Talukdar, Fenna M. Krienen, Hesheng Liu, Trey Hedden, Jessica R. Andrews-Hanna, Reisa A. Sperling, and Keith A. Johnson. Cortical hubs revealed by intrinsic functional connectivity: mapping, assessment of stability, and relation to Alzheimer’s disease. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 29(6):1860–1873, February 2009. ISSN 1529-2401. doi: 10.1523/JNEUROSCI.5062-08.2009. URL <http://dx.doi.org/10.1523/JNEUROSCI.5062-08.2009>.

F. Carbonell, P. Bellec, and A. Shmuel. Quantification of the impact of a confounding variable on functional connectivity confirms anti-correlated networks in the resting-state. *Neuroimage*, 86:343–353, Feb 2014. doi: 10.1016/j.neuroimage.2013.10.013. URL <http://dx.doi.org/10.1016/j.neuroimage.2013.10.013>.

Gang Chen, B. Douglas Ward, Chunming Xie, Wenjun Li, Zhilin Wu, Jennifer L. Jones, Małgorzata Franczak, Piero Antuono, and Shi-Jiang Li. Classification of Alzheimer Disease, Mild Cognitive Impairment, and Normal Cognitive Status with Large-Scale Network Analysis Based on Resting-State Functional MR Imaging. *Radiology*, 259

- (1):213–221, April 2011. doi: 10.1148/radiol.10100734. URL <http://dx.doi.org/10.1148/radiol.10100734>.
- G. Chetelat, B. Desgranges, V. de la Sayette, F. Viader, F. Eustache, and J-C. Baron. Mild cognitive impairment: Can fdg-pet predict who is to rapidly convert to alzheimer’s disease? *Neurology*, 60(8):1374–1377, Apr 2003.
- Z. H. Cho, S. C. Chung, D. W. Lim, and E. K. Wong. Effects of the acoustic noise of the gradient systems on fmri: a study on auditory, motor, and visual cortices. *Magn Reson Med*, 39(2):331–335, Feb 1998.
- D. L. Collins, A. P. Zijdenbos, V. Kollokian, J. G. Sled, N. J. Kabani, C. J. Holmes, and A. C. Evans. Design and construction of a realistic digital brain phantom. *IEEE Trans Med Imaging*, 17(3):463–468, June 1998. ISSN 0278-0062. URL <http://view.ncbi.nlm.nih.gov/pubmed/9735909>.
- Dietmar Cordes, Victor M. Haughton, Konstantinos Arfanakis, Gary J. Wendt, Patrick A. Turski, Chad H. Moritz, Michelle A. Quigley, and M. Elizabeth Meyerand. Mapping Functionally Related Regions of Brain with Functional Connectivity MR Imaging. *American Journal of Neuroradiology*, 21(9):1636–1644, October 2000. URL <http://www.ajnr.org/content/21/9/1636.abstract>.
- Dietmar Cordes, Victor M. Haughton, Konstantinos Arfanakis, John D. Carew, Patrick A. Turski, Chad H. Moritz, Michelle A. Quigley, and Elizabeth M. Meyerand. Frequencies Contributing to Functional Connectivity in the Cerebral Cortex in "Resting-state" Data. *AJNR Am J Neuroradiol*, 22(7):1326–1333, August 2001. URL <http://www.ajnr.org/cgi/content/abstract/22/7/1326>.
- Dietmar Cordes, Vic Haughton, John D. Carew, Konstantinos Arfanakis, and Ken Maravilla. Hierarchical clustering to measure connectivity in fMRI resting-state data. *Magnetic Resonance Imaging*, 20(4):305–317, May 2002. ISSN 0730725X. doi: 10.1016/S0730-725X(02)00503-9. URL [http://dx.doi.org/10.1016/S0730-725X\(02\)00503-9](http://dx.doi.org/10.1016/S0730-725X(02)00503-9).

Corinna Cortes and Vladimir Vapnik. Support-vector networks. *Machine Learning*, 20(3):273–297, September 1995. ISSN 0885-6125. doi: 10.1007/BF00994018. URL <http://dx.doi.org/10.1007/BF00994018>.

Sergi G Costafreda, Akash Khanna, Janaina Mourao-Miranda, and Cynthia H Y Fu. Neural correlates of sad faces predict clinical remission to cognitive behavioural therapy in depression. *Neuroreport*, 20(7):637–641, May 2009. doi: 10.1097/WNR.0b013e3283294159. URL <http://dx.doi.org/10.1097/WNR.0b013e3283294159>.

Kathryn R Cullen, Dylan G Gee, Bonnie Klimes-Dougan, Vilma Gabbay, Leslie Hulvershorn, Bryon A Mueller, Jazmin Camchong, Christopher J Bell, Alaa Houri, Sanjiv Kumra, Kelvin O Lim, F. Xavier Castellanos, and Michael P Milham. A preliminary study of functional connectivity in comorbid adolescent depression. *Neurosci Lett*, 460(3):227–231, Sep 2009. doi: 10.1016/j.neulet.2009.05.022. URL <http://dx.doi.org/10.1016/j.neulet.2009.05.022>.

Zhengjia Dai, Chaogan Yan, Kuncheng Li, Zhiqun Wang, Jinhui Wang, Miao Cao, Qixiang Lin, Ni Shu, Mingrui Xia, Yanchao Bi, and Yong He. Identifying and mapping connectivity patterns of brain network hubs in alzheimer’s disease. *Cereb Cortex*, Oct 2014. doi: 10.1093/cercor/bhu246. URL <http://dx.doi.org/10.1093/cercor/bhu246>.

Marcello D’Amelio and Paolo Maria Rossini. Brain excitability and connectivity of neuronal assemblies in alzheimer’s disease: from animal models to human findings. *Progress in neurobiology*, 99(1):42–60, October 2012. ISSN 1873-5118. URL <http://view.ncbi.nlm.nih.gov/pubmed/22789698>.

J. S. Damoiseaux, S. A. R. B. Rombouts, F. Barkhof, P. Scheltens, C. J. Stam, S. M. Smith, and C. F. Beckmann. Consistent resting-state networks across healthy subjects. *Proceedings of the National Academy of Sciences*, 103(37):13848–13853, September 2006. ISSN 1091-6490. doi: 10.1073/pnas.0601417103. URL <http://dx.doi.org/10.1073/pnas.0601417103>.

Jessica S. Damoiseaux, Katherine E. Prater, Bruce L. Miller, and Michael D. Grecius. Functional connectivity tracks clinical deterioration in Alzheimer's disease. *Neurobiology of Aging*, 33(4):828.e19–828.e30, April 2012. ISSN 01974580. doi: 10.1016/j.neurobiolaging.2011.06.024. URL <http://dx.doi.org/10.1016/j.neurobiolaging.2011.06.024>.

Christian Dansereau, Celine Risterucci, Emilio Merlo Pich, Douglas Arnold, and Pierre Bellec. A power analysis for multisite studies in resting-state functional connectivity, with an application to clinical trials in alzheimer's disease. volume 9, pages P248 – P249, 2013. doi: <http://dx.doi.org/10.1016/j.jalz.2013.05.489>. URL <http://www.sciencedirect.com/science/article/pii/S1552526013011461>. Alzheimer's Association International Conference 2013 Alzheimer's Association International Conference 2013.

Christian Dansereau, Pierre Bellec, Kangjoo Lee, Francesca Pittau, Jean Gotman, and Christophe Grova. Detection of abnormal resting-state networks in individual patients suffering from focal epilepsy: An initial step toward individual connectivity assessment. *Frontiers in Neuroscience*, 8(419), 2014a. ISSN 1662-453X. doi: 10.3389/fnins.2014.00419. URL http://www.frontiersin.org/brain_imaging_methods/10.3389/fnins.2014.00419/abstract.

Christian Dansereau, Pierre Orban, and Pierre Bellec. Impact of motion on resting-state fmri connectivity in healthy aging and alzheimer's disease, and possible remedies. volume 10, pages P49 –, 2014b. doi: <http://dx.doi.org/10.1016/j.jalz.2014.05.091>. URL <http://www.sciencedirect.com/science/article/pii/S1552526014007389>. Alzheimer's Association International Conference 2014 Alzheimer's Association International Conference 2014.

John Desmond and Gary Glover. Estimating sample size in functional mri (fmri) neuroimaging studies: Statistical power analyses. *Journal of Neuroscience Methods*, 118 (2):115–128, August 2002. ISSN 01650270. URL [http://dx.doi.org/10.1016/s0165-0270\(02\)00121-8](http://dx.doi.org/10.1016/s0165-0270(02)00121-8).

- V. Edward, C. Windischberger, R. Cunnington, M. Erdler, R. Lanzenberger, D. Mayer, W. Endl, and R. Beisteiner. Quantification of fmri artifact reduction by a novel plaster cast head holder. *Hum Brain Mapp*, 11(3):207–213, Nov 2000.
- Bradley Efron and R. J. Tibshirani. *An Introduction to the Bootstrap (Chapman & Hall/CRC Monographs on Statistics & Applied Probability)*. Chapman and Hall/CRC, 1 edition, May 1994. ISBN 0412042312. URL <http://www.worldcat.org/isbn/0412042312>.
- M. R. Elliott, R. W. Bowtell, and P. G. Morris. The effect of scanner sound in visual, motor, and auditory functional mri. *Magn Reson Med*, 41(6):1230–1235, Jun 1999.
- Damien A Fair, Joel T Nigg, Swathi Iyer, Deepti Bathula, Kathryn L Mills, Nico U F Dosenbach, Bradley L Schlaggar, Maarten Mennes, David Gutman, Saroja Bangaru, Jan K Buitelaar, Daniel P Dickstein, Adriana Di Martino, David N Kennedy, Clare Kelly, Beatriz Luna, Julie B Schweitzer, Katerina Velanova, Yu-Feng Wang, Stewart Mostofsky, F. Xavier Castellanos, and Michael P Milham. Distinct neural signatures detected for adhd subtypes after controlling for micro-movements in resting state functional connectivity mri data. *Front Syst Neurosci*, 6:80, 2012. doi: 10.3389/fnsys.2012.00080. URL <http://dx.doi.org/10.3389/fnsys.2012.00080>.
- Yong Fan, Dinggang Shen, and Christos Davatzikos. Classification of structural images via high-dimensional image warping, robust feature extraction, and svm. *Med Image Comput Comput Assist Interv*, 8(Pt 1):1–8, 2005.
- Fang Fang, Scott O Murray, and Sheng He. Duration-dependent fmri adaptation and distributed viewer-centered face representation in human visual cortex. *Cereb Cortex*, 17(6):1402–1411, Jun 2007. doi: 10.1093/cercor/bhl053. URL <http://dx.doi.org/10.1093/cercor/bhl053>.
- David A Feinberg, Steen Moeller, Stephen M Smith, Edward Auerbach, Sudhir Ramanan, Matthias Gunther, Matt F Glasser, Karla L Miller, Kamil Ugurbil, and Essa Yacoub. Multiplexed echo planar imaging for sub-second whole brain fmri and

- fast diffusion imaging. *PLoS One*, 5(12):e15710, 2010. doi: 10.1371/journal.pone.0015710. URL <http://dx.doi.org/10.1371/journal.pone.0015710>.
- Vladimir Fonov, Alan C. Evans, Kelly Botteron, C. Robert Almli, Robert C. McKinstry, D. Louis Collins, and Brain Development Cooperative Group. Unbiased average age-appropriate atlases for pediatric studies. *NeuroImage*, 54(1):313–327, January 2011. ISSN 1095-9572. doi: 10.1016/j.neuroimage.2010.07.033. URL <http://dx.doi.org/10.1016/j.neuroimage.2010.07.033>.
- Michael D. Fox and Marcus E. Raichle. Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Nat Rev Neurosci*, 8(9):700–711, September 2007. ISSN 1471-003X. doi: 10.1038/nrn2201. URL <http://dx.doi.org/10.1038/nrn2201>.
- Michael D Fox, Dongyang Zhang, Abraham Z Snyder, and Marcus E Raichle. The global signal and observed anticorrelated resting state brain networks. *J Neurophysiol*, 101(6):3270–3283, Jun 2009. doi: 10.1152/jn.90777.2008. URL <http://dx.doi.org/10.1152/jn.90777.2008>.
- Yoav Freund and Robert E Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences*, 55(1):119 – 139, 1997. ISSN 0022-0000. doi: <http://dx.doi.org/10.1006/jcss.1997.1504>. URL <http://www.sciencedirect.com/science/article/pii/S00220009791504X>.
- Lee Friedman and Gary Glover. Report on a multicenter fmri quality assurance protocol. *Journal of magnetic resonance imaging : JMRI*, 23(6):827–839, June 2006. ISSN 1053-1807. URL <http://dx.doi.org/10.1002/jmri.20583>.
- Lee Friedman, Gary Glover, and The FBIRN Consortium. Reducing interscanner variability of activation in a multicenter fmri study: Controlling for signal-to-fluctuation-noise-ratio (sfnr) differences. *NeuroImage*, 33(2):471–481, November 2006. ISSN

10538119. URL <http://dx.doi.org/10.1016/j.neuroimage.2006.07.012>.

Cynthia H Y Fu, Janaina Mourao-Miranda, Sergi G Costafreda, Akash Khanna, Andre F Marquand, Steve C R Williams, and Michael J Brammer. Pattern classification of sad facial processing: toward the development of neurobiological markers in depression. *Biol Psychiatry*, 63(7):656–662, Apr 2008. doi: 10.1016/j.biopsych.2007.08.020. URL <http://dx.doi.org/10.1016/j.biopsych.2007.08.020>.

Ran Gilad-Bachrach, Amir Navot, and Naftali Tishby. Margin based feature selection - theory and algorithms. In *Proceedings of the Twenty-first International Conference on Machine Learning*, ICML '04, pages 43–, New York, NY, USA, 2004. ACM. ISBN 1-58113-838-5. doi: 10.1145/1015330.1015352. URL <http://doi.acm.org/10.1145/1015330.1015352>.

Federico Giove, Tommaso Gili, Vittorio Iacobella, Emiliano Macaluso, and Bruno Martaviglia. Images-based suppression of unwanted global signals in resting-state functional connectivity studies. *Magnetic resonance imaging*, 27(8):1058–1064, October 2009. ISSN 1873-5894. doi: 10.1016/j.mri.2009.06.004. URL <http://dx.doi.org/10.1016/j.mri.2009.06.004>.

Joseph S. Goveas, Chunming Xie, B. Douglas Ward, Zhilin Wu, Wenjun Li, Małgorzata Franczak, Jennifer L. Jones, Piero G. Antuono, and Shi-Jiang J. Li. Recovery of hippocampal network connectivity correlates with cognitive improvement in mild Alzheimer's disease patients treated with donepezil assessed by resting-state fMRI. *Journal of magnetic resonance imaging : JMRI*, 34(4):764–773, October 2011. ISSN 1522-2586. doi: 10.1002/jmri.22662. URL <http://dx.doi.org/10.1002/jmri.22662>.

M. D. Greicius, B. Krasnow, A. L. Reiss, and V. Menon. Functional connectivity in the resting brain: A network analysis of the default mode hypothesis. *Proc. Natl. Acad. Sci.*, 100(1):253–258, January 2003. ISSN 1091-6490. doi: 10.1073/pnas.0135058100. URL <http://dx.doi.org/10.1073/pnas.0135058100>.

Michael D. Greicius, Gaurav Srivastava, Allan L. Reiss, and Vinod Menon. Default-mode network activity distinguishes Alzheimer's disease from healthy aging: Evidence from functional MRI. *Proceedings of the National Academy of Sciences of the United States of America*, 101(13):4637–4642, March 2004. ISSN 1091-6490. doi: 10.1073/pnas.0308627101. URL <http://dx.doi.org/10.1073/pnas.0308627101>.

Michael D. Greicius, Benjamin H. Flores, Vinod Menon, Gary H. Glover, Hugh B. Solvason, Heather Kenna, Allan L. Reiss, and Alan F. Schatzberg. Resting-state functional connectivity in major depression: abnormally increased contributions from subgenual cingulate cortex and thalamus. *Biological psychiatry*, 62(5):429–437, September 2007. ISSN 0006-3223. doi: 10.1016/j.biopsych.2006.09.020. URL <http://dx.doi.org/10.1016/j.biopsych.2006.09.020>.

Michael D. Greicius, Kaustubh Supekar, Vinod Menon, and Robert F. Dougherty. Resting-state functional connectivity reflects structural connectivity in the default mode network. *Cerebral cortex (New York, N.Y. : 1991)*, 19(1):72–78, January 2009. ISSN 1460-2199. doi: 10.1093/cercor/bhn059. URL <http://dx.doi.org/10.1093/cercor/bhn059>.

D. A. Gusnard and M. E. Raichle. Searching for a baseline: Functional imaging and the resting human brain. *Nature Reviews Neuroscience*, 2(10):685–694, October 2001. ISSN 1471-003X. doi: 10.1038/35094500. URL <http://dx.doi.org/10.1038/35094500>.

Tim Hahn, Andre F Marquand, Ann-Christine Ehli, Thomas Dresler, Sarah Kittel-Schneider, Tomasz A Jarczok, Klaus-Peter Lesch, Peter M Jakob, Janaina Mourao-Miranda, Michael J Brammer, and Andreas J Fallgatter. Integrating neurobiological markers of depression. *Arch Gen Psychiatry*, 68(4):361–368, Apr 2011. doi: 10.1001/archgenpsychiatry.2010.178. URL <http://dx.doi.org/10.1001/archgenpsychiatry.2010.178>.

M. Hampson, B. S. Peterson, P. Skudlarski, J. C. Gatenby, and J. C. Gore. Detection of

- functional connectivity using temporal correlations in MR images. *Hum Brain Mapp*, 15(4):247–262, April 2002. ISSN 1065-9471. URL <http://view.ncbi.nlm.nih.gov/pubmed/11835612>.
- John Hardy and Dennis J Selkoe. The amyloid hypothesis of alzheimer’s disease: progress and problems on the road to therapeutics. *Science*, 297(5580):353–356, Jul 2002. doi: 10.1126/science.1072994. URL <http://dx.doi.org/10.1126/science.1072994>.
- Egbert Hartstra, Simone KÃijhn, Tom Verguts, and Marcel Brass. The implementation of verbal instructions: an fmri study. *Hum Brain Mapp*, 32(11):1811–1824, Nov 2011. doi: 10.1002/hbm.21152. URL <http://dx.doi.org/10.1002/hbm.21152>.
- Yong He, Zhang Chen, Gaolang Gong, and Alan Evans. Neuronal networks in Alzheimer’s disease. *The Neuroscientist : a review journal bringing neurobiology, neurology and psychiatry*, 15(4):333–350, August 2009. ISSN 1073-8584. doi: 10.1177/1073858409334423. URL <http://dx.doi.org/10.1177/1073858409334423>.
- David J Heeger and David Ress. What does fmri tell us about neuronal activity? *Nat Rev Neurosci*, 3(2):142–151, Feb 2002. doi: 10.1038/nrn730. URL <http://dx.doi.org/10.1038/nrn730>.
- R. D. Hoge, J. Atkinson, B. Gill, G. R. Crelier, S. Marrett, and G. B. Pike. Investigation of bold signal dependence on cerebral blood flow and oxygen consumption: the deoxyhemoglobin dilution model. *Magn Reson Med*, 42(5):849–863, Nov 1999.
- James X. Hu, Hongyu Zhao, and Harrison H. Zhou. False Discovery Rate Control With Groups. *Journal of the American Statistical Association*, 105(491):1215–1227, September 2010. ISSN 0162-1459. doi: 10.1198/jasa.2010.tm09329. URL <http://dx.doi.org/10.1198/jasa.2010.tm09329>.
- C. R. Jack, R. C. Petersen, Y. C. Xu, P. C. O’Brien, G. E. Smith, R. J. Ivnik, B. F. Boeve, S. C. Waring, E. G. Tangalos, and E. Kokmen. Prediction of ad with mri-based

hippocampal volume in mild cognitive impairment. *Neurology*, 52(7):1397–1403, Apr 1999.

Clifford R Jack, David S Knopman, William J Jagust, Leslie M Shaw, Paul S Aisen, Michael W Weiner, Ronald C Petersen, and John Q Trojanowski. Hypothetical model of dynamic biomarkers of the alzheimer’s pathological cascade. *Lancet Neurol*, 9(1):119–128, Jan 2010. doi: 10.1016/S1474-4422(09)70299-6. URL [http://dx.doi.org/10.1016/S1474-4422\(09\)70299-6](http://dx.doi.org/10.1016/S1474-4422(09)70299-6).

Jaeseung Jeong. Eeg dynamics in patients with alzheimer’s disease. *Clin Neurophysiol*, 115(7):1490–1505, Jul 2004. doi: 10.1016/j.clinph.2004.01.001. URL <http://dx.doi.org/10.1016/j.clinph.2004.01.001>.

Tianzi Jiang, Yong He, Yufeng Zang, and Xuchu Weng. Modulation of functional connectivity during the resting state and the motor task. *Hum. Brain Mapp.*, 22(1):63–71, May 2004. ISSN 1065-9471. doi: 10.1002/hbm.20012. URL <http://dx.doi.org/10.1002/hbm.20012>.

Xi Jiang, Xin Zhang, and Dajiang Zhu. Intrinsic functional component analysis via sparse representation on alzheimer’s disease neuroimaging initiative database. *Brain Connect*, 4(8):575–586, Oct 2014. doi: 10.1089/brain.2013.0221. URL <http://dx.doi.org/10.1089/brain.2013.0221>.

Yasuhiro Kawasaki, Michio Suzuki, Ferath Kherif, Tsutomu Takahashi, Shi-Yu Zhou, Kazue Nakamura, Mie Matsui, Tomiki Sumiyoshi, Hikaru Seto, and Masayoshi Kurachi. Multivariate voxel-based morphometry successfully differentiates schizophrenia patients from healthy controls. *Neuroimage*, 34(1):235–242, Jan 2007. doi: 10.1016/j.neuroimage.2006.08.018. URL <http://dx.doi.org/10.1016/j.neuroimage.2006.08.018>.

Clare Kelly, Roberto Toro, Adriana Di Martino, Christine L. Cox, Pierre Bellec, F. Xavier Castellanos, and Michael P. Milham. A convergent functional architecture of the insula emerges across imaging modalities. *NeuroImage*, 61(4):1129–

- 1142, July 2012. ISSN 10538119. doi: 10.1016/j.neuroimage.2012.03.021. URL <http://dx.doi.org/10.1016/j.neuroimage.2012.03.021>.
- Markus Klarhofer, Markus Barth, and Ewald Moser. Comparison of multi-echo spiral and echo planar imaging in functional mri. *Magn Reson Imaging*, 20(4):359–364, May 2002.
- Igor Kononenko, Edvard Simec, and Marko Robnik-Sikonja. Overcoming the myopia of inductive learning algorithms with relieff. *Applied Intelligence*, 7:39–55, 1997.
- S. B. Kotsiantis. Supervised machine learning: A review of classification techniques. pages 3–24, 2007. URL <http://dl.acm.org/citation.cfm?id=1566770.1566773>.
- Zhiqiang Lao, Dinggang Shen, Zhong Xue, Bilge Karacali, Susan M Resnick, and Christos Davatzikos. Morphological classification of brains via high-dimensional shape transformations and machine learning methods. *Neuroimage*, 21(1):46–57, Jan 2004.
- Meng Liang, Yuan Zhou, Tianzi Jiang, Zhening Liu, Lixia Tian, Haihong Liu, and Yihui Hao. Widespread functional connectivity in schizophrenia with resting-state functional magnetic resonance imaging. *Neuroreport*, 17(2):209–213, Feb 2006.
- Fa-Hsuan Lin, Teng-Yi Huang, Nan-Kuei Chen, Fu-Nien Wang, Steven M Stufflebeam, John W Belliveau, Lawrence L Wald, and Kenneth K Kwong. Functional mri using regularized parallel imaging acquisition. *Magn Reson Med*, 54(2):343–353, Aug 2005. doi: 10.1002/mrm.20555. URL <http://dx.doi.org/10.1002/mrm.20555>.
- Yong Liu, Chunshui Yu, Meng Liang, Jun Li, Lixia Tian, Yuan Zhou, Wen Qin, Kuncheng Li, and Tianzi Jiang. Whole brain functional connectivity in the early blind. *Brain*, 130(8):2085–2096, August 2007. ISSN 1460-2156. doi: 10.1093/brain/awm121. URL <http://dx.doi.org/10.1093/brain/awm121>.
- Mark J Lowe, Micheal D Phillips, Joseph T Lurito, David Mattson, Mario Dzemidzic, and Vincent P Mathews. Multiple sclerosis: low-frequency temporal blood oxygen

- level-dependent fluctuations indicate reduced functional connectivity initial results. *Radiology*, 224(1):184–192, Jul 2002. doi: 10.1148/radiol.2241011005. URL <http://dx.doi.org/10.1148/radiol.2241011005>.
- MJ Lowe, BJ Mock, and JA Sorenson. Functional connectivity in single and multislice echoplanar imaging using resting-state fluctuations. *NeuroImage*, 7(2):119–132, February 1998. ISSN 1053-8119. URL <http://dx.doi.org/10.1006/nimg.1997.0315>.
- Torben E. Lund, Kristoffer H. Madsen, Karam Sidaros, Wen-Lin Luo, and Thomas E. Nichols. Non-white noise in fMRI: does modelling have an impact? *NeuroImage*, 29(1):54–66, January 2006. ISSN 1053-8119. doi: 10.1016/j.neuroimage.2005.07.005. URL <http://dx.doi.org/10.1016/j.neuroimage.2005.07.005>.
- Daniel S. Margulies, Justin L. Vincent, Clare Kelly, Gabriele Lohmann, Lucina Q. Uddin, Bharat B. Biswal, Arno Villringer, F. Xavier Castellanos, Michael P. Milham, and Michael Petrides. Precuneus shares intrinsic functional architecture in humans and monkeys. *Proceedings of the National Academy of Sciences of the United States of America*, November 2009. ISSN 1091-6490. doi: 10.1073/pnas.0905314106. URL <http://dx.doi.org/10.1073/pnas.0905314106>.
- Andre F Marquand, Janaina Mourão-Miranda, Michael J Brammer, Anthony J Cleare, and Cynthia H Y Fu. Neuroanatomy of verbal working memory as a diagnostic biomarker for depression. *Neuroreport*, 19(15):1507–1511, Oct 2008. doi: 10.1097/WNR.0b013e328310425e. URL <http://dx.doi.org/10.1097/WNR.0b013e328310425e>.
- Niklas Mattsson, Henrik Zetterberg, Oskar Hansson, Niels Andreasen, Lucilla Parnetti, Michael Jonsson, Sanna-Kaisa Herukka, Wiesje M van der Flier, Marinus A Blankenstein, Michael Ewers, Kenneth Rich, Elmar Kaiser, Marcel Verbeek, Magda Tsolaki, Ezra Mulugeta, Erik Rosén, Dag Aarsland, Pieter Jelle Visser, Johannes Schröder, Jan Marcusson, Mony de Leon, Harald Hampel, Philip Scheltens, Tuula Pirttilä, Anders Wallin, Maria Eriksdotter Jähnigen, Lennart Minthon, Bengt Winblad, and

- Kaj Blennow. Csf biomarkers and incipient alzheimer disease in patients with mild cognitive impairment. *JAMA*, 302(4):385–393, Jul 2009. doi: 10.1001/jama.2009.1064. URL <http://dx.doi.org/10.1001/jama.2009.1064>.
- M. J. McKeown, S. Makeig, G. G. Brown, T. P. Jung, S. S. Kindermann, A. J. Bell, and T. J. Sejnowski. Analysis of fMRI data by blind separation into independent spatial components. *Hum Brain Mapp*, 6(3):160–188, 1998. ISSN 1065-9471. URL <http://view.ncbi.nlm.nih.gov/pubmed/9673671>.
- G. McKhann, D. Drachman, M. Folstein, R. Katzman, D. Price, and E. M. Stadlan. 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(7):939–944, July 1984. ISSN 0028-3878. URL <http://www.neurology.org/content/34/7/939.abstract>.
- V. Menon, K.O. Lim, J.H. Anderson, J. Johnson, and A. Pfefferbaum. Design and efficacy of a head-coil bite bar for reducing movement-related artifacts during functional mri scanning. *Behavior Research Methods, Instruments, & Computers*, 29(4):589–594, 1997. ISSN 0743-3808. doi: 10.3758/BF03210613. URL <http://dx.doi.org/10.3758/BF03210613>.
- A. Meyer-Baese, A. Wismueller, and O. Lange. Comparison of two exploratory data analysis methods for fMRI: unsupervised clustering versus independent component analysis. *IEEE Trans Inf Technol Biomed*, 8(3):387–398, September 2004. ISSN 1089-7771. URL <http://view.ncbi.nlm.nih.gov/pubmed/15484444>.
- Michal Mikl, Radek Marecek, Petr Hlustík, Martina Pavlicová, Ales Drastich, Pavel Chlebus, Milan Brázdil, and Petr Krupa. Effects of spatial smoothing on fmri group inferences. *Magnetic resonance imaging*, 26(4):490–503, May 2008. ISSN 0730-725X. URL <http://view.ncbi.nlm.nih.gov/pubmed/18060720>.
- John C Morris. Early-stage and preclinical alzheimer disease. *Alzheimer Dis Assoc Disord*, 19(3):163–165, 2005.

- Janaina Mourao-Miranda, Arun L W Bokde, Christine Born, Harald Hampel, and Martin Stetter. Classifying brain states and determining the discriminating activation patterns: Support vector machine on functional mri data. *Neuroimage*, 28(4):980–995, Dec 2005. doi: 10.1016/j.neuroimage.2005.06.070. URL <http://dx.doi.org/10.1016/j.neuroimage.2005.06.070>.
- Susanne G Mueller, Michael W Weiner, Leon J Thal, Ronald C Petersen, Clifford Jack, William Jagust, John Q Trojanowski, Arthur W Toga, and Laurel Beckett. The alzheimer’s disease neuroimaging initiative. *Neuroimaging Clin N Am*, 15(4):869–77, xi–xii, Nov 2005. doi: 10.1016/j.nic.2005.09.008. URL <http://dx.doi.org/10.1016/j.nic.2005.09.008>.
- Kevin Murphy, Rasmus M. Birn, Daniel A. Handwerker, Tyler B. Jones, and Peter A. Bandettini. The impact of global signal regression on resting state correlations: are anti-correlated networks introduced? *NeuroImage*, 44(3):893–905, February 2009. ISSN 1095-9572. doi: 10.1016/j.neuroimage.2008.09.036. URL <http://dx.doi.org/10.1016/j.neuroimage.2008.09.036>.
- Jared Nielsen, Brandon Zielinski, Thomas Fletcher, Andrew Alexander, Nicholas Lange, Erin Bigler, Janet Lainhart, and Jeffrey Anderson. Multisite functional connectivity mri classification of autism: Abide results. *Frontiers in human neuroscience*, 7:–, 2013. ISSN 1662-5161. URL <http://view.ncbi.nlm.nih.gov/pubmed/24093016>.
- Ilia Nouretdinov, Sergi Costafreda, Alexander Gammerman, Alexey Chervonenkis, Vladimir Vovk, Vladimir Vapnik, and Cynthia Fu. Machine learning classification with confidence: Application of transductive conformal predictors to mri-based diagnostic and prognostic markers in depression. *NeuroImage*, 56(2):809–813, May 2011. ISSN 10538119. URL <http://dx.doi.org/10.1016/j.neuroimage.2010.05.023>.
- S. Ogawa, T. M. Lee, A. R. Kay, and D. W. Tank. Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proceedings of the National Academy of*

- Sciences*, 87(24):9868–9872, December 1990. ISSN 1091-6490. doi: 10.1073/pnas.87.24.9868. URL <http://dx.doi.org/10.1073/pnas.87.24.9868>.
- R. C. Petersen. Mild cognitive impairment as a diagnostic entity. *J Intern Med*, 256(3):183–194, Sep 2004. doi: 10.1111/j.1365-2796.2004.01388.x. URL <http://dx.doi.org/10.1111/j.1365-2796.2004.01388.x>.
- Jonathan D. Power, Alexander L. Cohen, Steven M. Nelson, Gagan S. Wig, Kelly A. Barnes, Jessica A. Church, Alecia C. Vogel, Timothy O. Laumann, Fran M. Miezin, Bradley L. Schlaggar, and Steven E. Petersen. Functional Network Organization of the Human Brain. *Neuron*, 72(4):665–678, November 2011. ISSN 08966273. doi: 10.1016/j.neuron.2011.09.006. URL <http://dx.doi.org/10.1016/j.neuron.2011.09.006>.
- Jonathan D. Power, Kelly A. Barnes, Abraham Z. Snyder, Bradley L. Schlaggar, and Steven E. Petersen. Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *NeuroImage*, 59(3):2142–2154, February 2012. ISSN 1095-9572. doi: 10.1016/j.neuroimage.2011.10.018. URL <http://dx.doi.org/10.1016/j.neuroimage.2011.10.018>.
- Jonathan D. Power, Anish Mitra, Timothy O. Laumann, Abraham Z. Snyder, Bradley L. Schlaggar, and Steven E. Petersen. Methods to detect, characterize, and remove motion artifact in resting state fMRI. *NeuroImage*, 84:320–341, January 2014. doi: 10.1016/j.neuroimage.2013.08.048. URL <http://dx.doi.org/10.1016/j.neuroimage.2013.08.048>.
- Chris Rorden, Hans-Otto Karnath, and Leonardo Bonilha. Improving lesion-symptom mapping. *J Cogn Neurosci*, 19(7):1081–1088, Jul 2007. doi: 10.1162/jocn.2007.19.7.1081. URL <http://dx.doi.org/10.1162/jocn.2007.19.7.1081>.
- Ziad S. Saad, Stephen J. Gotts, Kevin Murphy, Gang Chen, Hang Joon J. Jo, Alex Martin, and Robert W. Cox. Trouble at rest: how correlation patterns and group differences become distorted after global signal regression. *Brain connectivity*, 2(1):25–32, 2012.

- ISSN 2158-0022. doi: 10.1089/brain.2012.0080. URL <http://dx.doi.org/10.1089/brain.2012.0080>.
- R. Salvador, J. Suckling, M. R. Coleman, J. D. Pickard, D. Menon, and E. Bullmore. Neurophysiological architecture of functional magnetic resonance images of human brain. *Cereb Cortex*, 15(9):1332–1342, September 2005. ISSN 1047-3211. URL <http://view.ncbi.nlm.nih.gov/pubmed/15635061>.
- R. Salvador, A. Martínez, E. Pomarol-Clotet, S. Sarró, J. Suckling, and E. Bullmore. Frequency based mutual information measures between clusters of brain regions in functional magnetic resonance imaging. *NeuroImage*, 35(1):83–88, March 2007. ISSN 1053-8119. doi: 10.1016/j.neuroimage.2006.12.001. URL <http://dx.doi.org/10.1016/j.neuroimage.2006.12.001>.
- Robert E. Schapire, Yoav Freund, Peter Bartlett, and Wee Sun Lee. Boosting the margin: a new explanation for the effectiveness of voting methods. *The Annals of Statistics*, 26(5):1651–1686, 10 1998. doi: 10.1214/aos/1024691352. URL <http://dx.doi.org/10.1214/aos/1024691352>.
- Ganesh M Shankar, Shaomin Li, Tapan H Mehta, Amaya Garcia-Munoz, Nina E Shepardson, Imelda Smith, Francesca M Brett, Michael A Farrell, Michael J Rowan, Cynthia A Lemere, Ciaran M Regan, Dominic M Walsh, Bernardo L Sabatini, and Dennis J Selkoe. Amyloid-beta protein dimers isolated directly from alzheimer's brains impair synaptic plasticity and memory. *Nat Med*, 14(8):837–842, Aug 2008. doi: 10.1038/nm1782. URL <http://dx.doi.org/10.1038/nm1782>.
- Zarrar Shehzad, Clare M. Kelly, Philip T. Reiss, Dylan G. Gee, Kristin Gotimer, Lucina Q. Uddin, Sang Han H. Lee, Daniel S. Margulies, Amy Krain K. Roy, Bharat B. Biswal, Eva Petkova, F. Xavier Castellanos, and Michael P. Milham. The resting brain: unconstrained yet reliable. *Cerebral cortex (New York, N.Y. : 1991)*, 19(10):2209–2229, October 2009. ISSN 1460-2199. doi: 10.1093/cercor/bhn256. URL <http://dx.doi.org/10.1093/cercor/bhn256>.

Yvette Sheline and Marcus Raichle. Resting state functional connectivity in pre-clinical alzheimerâŽs disease. *Biological Psychiatry*, 74(5):340â€“347, September 2013. ISSN 00063223. URL <http://dx.doi.org/10.1016/j.biopsych.2012.11.028>.

Reisa A. Sperling, Paul S. Aisen, Laurel A. Beckett, David A. Bennett, Suzanne Craft, Anne M. Fagan, Takeshi Iwatsubo, Clifford R. Jack, Jeffrey Kaye, Thomas J. Montine, Denise C. Park, Eric M. Reiman, Christopher C. Rowe, Eric Siemers, Yaakov Stern, Kristine Yaffe, Maria C. Carrillo, Bill Thies, Marcelle Morrison-Bogorad, Molly V. Wagster, and Creighton H. Phelps. Toward defining the preclinical stages of Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's & Dementia*, 7(3):280â€“292, May 2011. doi: 10.1016/j.jalz.2011.03.003. URL <http://dx.doi.org/10.1016/j.jalz.2011.03.003>.

Lixia Tian, Tianzi Jiang, Meng Liang, Xiaobo Li, Yong He, Kun Wang, Bingli Cao, and Tao Jiang. Stabilities of negative correlations between blood oxygen level-dependent signals associated with sensory and motor cortices. *Hum Brain Mapp*, 28(7):681â€“690, Jul 2007. doi: 10.1002/hbm.20300. URL <http://dx.doi.org/10.1002/hbm.20300>.

Dardo Tomasi and Nora D. Volkow. Functional connectivity density mapping. *Proceedings of the National Academy of Sciences*, 107(21):9885â€“9890, May 2010. doi: 10.1073/pnas.1001414107. URL <http://dx.doi.org/10.1073/pnas.1001414107>.

N. Tzourio-Mazoyer, B. Landeau, D. Papathanassiou, F. Crivello, O. Etard, N. Delcroix, B. Mazoyer, and M. Joliot. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *NeuroImage*, 15(1):273â€“289, January 2002. ISSN 1053-8119. doi: 10.1006/nimg.2001.0978. URL <http://dx.doi.org/10.1006/nimg.2001.0978>.

Martijn van den Heuvel, Rene Mandl, and Hilleke Hulshoff Pol. Normalized Cut

- Group Clustering of Resting-State fMRI Data. *PLoS ONE*, 3(4):e2001+, April 2008. doi: 10.1371/journal.pone.0002001. URL <http://dx.doi.org/10.1371/journal.pone.0002001>.
- Martijn P. van den Heuvel, René C. Mandl, René S. Kahn, and Hilleke E. Hulshoff Pol. Functionally linked resting-state networks reflect the underlying structural connectivity architecture of the human brain. *Human brain mapping*, 30(10):3127–3141, October 2009. ISSN 1097-0193. doi: 10.1002/hbm.20737. URL <http://dx.doi.org/10.1002/hbm.20737>.
- Koene R. Van Dijk, Trey Hedden, Archana Venkataraman, Karleyton C. Evans, Sara W. Lazar, and Randy L. Buckner. Intrinsic functional connectivity as a tool for human connectomics: theory, properties, and optimization. *Journal of neurophysiology*, 103(1):297–321, January 2010. ISSN 1522-1598. doi: 10.1152/jn.00783.2009. URL <http://dx.doi.org/10.1152/jn.00783.2009>.
- G. Vanhoutte, M. Verhoye, and A. Van der Linden. Changing body temperature affects the $t2^*$ signal in the rat brain and reveals hypothalamic activity. *Magn Reson Med*, 55(5):1006–1012, May 2006. doi: 10.1002/mrm.20861. URL <http://dx.doi.org/10.1002/mrm.20861>.
- Archana Venkataraman, Marek Kubicki, Carl-Fredrik Westin, and Polina Golland. Robust feature selection in resting-state fmri connectivity based on population studies. *Conf Comput Vis Pattern Recognit Workshops*, pages 63–70, 2010. doi: 10.1109/CVPRW.2010.5543446. URL <http://dx.doi.org/10.1109/CVPRW.2010.5543446>.
- Anthony B. Waites, Regula S. Briellmann, Michael M. Saling, David F. Abbott, and Graeme D. Jackson. Functional connectivity networks are disrupted in left temporal lobe epilepsy. *Annals of neurology*, 59(2):335–343, February 2006. doi: 10.1002/ana.20733. URL <http://dx.doi.org/10.1002/ana.20733>.
- Kun Wang, Meng Liang, Liang Wang, Lixia Tian, Xinqing Zhang, Kuncheng Li, and

- Tianzi Jiang. Altered functional connectivity in early Alzheimer’s disease: a resting-state fMRI study. *Human brain mapping*, 28(10):967–978, October 2007. ISSN 1065-9471. doi: 10.1002/hbm.20324. URL <http://dx.doi.org/10.1002/hbm.20324>.
- Liang Wang, Yufeng Zang, Yong He, Meng Liang, Xinqing Zhang, Lixia Tian, Tao Wu, Tianzi Jiang, and Kuncheng Li. Changes in hippocampal connectivity in the early stages of Alzheimer’s disease: evidence from resting state fMRI. *NeuroImage*, 31(2):496–504, June 2006. ISSN 1053-8119. doi: 10.1016/j.neuroimage.2005.12.033. URL <http://dx.doi.org/10.1016/j.neuroimage.2005.12.033>.
- Cristen J. Willer, Yun Li, and Gonçalo R. Abecasis. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics*, 26(17):2190–2191, September 2010. doi: 10.1093/bioinformatics/btq340. URL <http://dx.doi.org/10.1093/bioinformatics/btq340>.
- A. M. Wink and J. B. T. M. Roerdink. BOLD noise assumptions in fMRI. *Int. J. Biomed. Imaging*, 2006, 2006. doi: 10.1155/IJBI/2006/12014. URL <http://dx.doi.org/10.1155/IJBI/2006/12014>.
- K. J. Worsley and K. J. Friston. Analysis of fMRI Time-Series Revisited—Again. *NeuroImage*, 2(3):173–181, September 1995. ISSN 10538119. doi: 10.1006/nimg.1995.1023. URL <http://dx.doi.org/10.1006/nimg.1995.1023>.
- Chao-Gan Yan, Brian Cheung, Clare Kelly, Stan Colcombe, R. Cameron Craddock, Adriana Di Martino, Qingyang Li, Xi-Nian Zuo, F. Xavier Castellanos, and Michael P. Milham. A comprehensive assessment of regional variation in the impact of head micromovements on functional connectomics. *NeuroImage*, 76:183–201, August 2013a. doi: 10.1016/j.neuroimage.2013.03.004. URL <http://dx.doi.org/10.1016/j.neuroimage.2013.03.004>.
- Chao-Gan G. Yan, Cameron C. Craddock, Xi-Nian N. Zuo, Yu-Feng F. Zang, and Michael P. Milham. Standardizing the intrinsic brain: towards robust measurement

- of inter-individual variation in 1000 functional connectomes. *NeuroImage*, 80:246–262, October 2013b. URL <http://view.ncbi.nlm.nih.gov/pubmed/23631983>.
- Chaogan Yan, Dongqiang Liu, Yong He, Qihong Zou, Chaozhe Zhu, Xinian Zuo, Xiangyu Long, and Yufeng Zang. Spontaneous Brain Activity in the Default Mode Network Is Sensitive to Different Resting-State Conditions with Limited Cognitive Load. *PLoS ONE*, 4(5):e5743+, May 2009. ISSN 1932-6203. doi: 10.1371/journal.pone.0005743. URL <http://dx.doi.org/10.1371/journal.pone.0005743>.
- Hong Yang, Xiang-Yu Long, Yihong Yang, Hao Yan, Chao-Zhe Zhu, Xiang-Ping Zhou, Yu-Feng Zang, and Qi-Yong Gong. Amplitude of low frequency fluctuation within visual areas revealed by resting-state functional mri. *Neuroimage*, 36(1):144–152, May 2007. doi: 10.1016/j.neuroimage.2007.01.054. URL <http://dx.doi.org/10.1016/j.neuroimage.2007.01.054>.
- Zhiwen Yu, Hau-San Wong, and Hongqiang Wang. Graph-based consensus clustering for class discovery from gene expression data. *Bioinformatics*, 23(21):2888–2896, November 2007. doi: 10.1093/bioinformatics/btm463. URL <http://dx.doi.org/10.1093/bioinformatics/btm463>.
- Y. Yuan, Z-X. Gu, and W-S. Wei. Fluorodeoxyglucose-positron-emission tomography, single-photon emission tomography, and structural mr imaging for prediction of rapid conversion to alzheimer disease in patients with mild cognitive impairment: a meta-analysis. *AJNR Am J Neuroradiol*, 30(2):404–410, Feb 2009. doi: 10.3174/ajnr.A1357. URL <http://dx.doi.org/10.3174/ajnr.A1357>.
- Yu-Feng F. Zang, Yong He, Chao-Zhe Z. Zhu, Qing-Jiu J. Cao, Man-Qiu Q. Sui, Meng Liang, Li-Xia X. Tian, Tian-Zi Z. Jiang, and Yu-Feng F. Wang. Altered baseline brain activity in children with ADHD revealed by resting-state functional MRI. *Brain & development*, 29(2):83–91, March 2007. ISSN 0387-7604. doi: 10.1016/j.braindev.2006.07.002. URL <http://dx.doi.org/10.1016/j.braindev.2006.07.002>.

- H. Zhang, S. Wang, J. Xing, B. Liu, Z. Ma, M. Yang, Z. Zhang, and G. Teng. Detection of PCC functional connectivity characteristics in resting-state fMRI in mild Alzheimer's disease. *Behavioural Brain Research*, 197(1):103–108, January 2009. ISSN 01664328. doi: 10.1016/j.bbr.2008.08.012. URL <http://dx.doi.org/10.1016/j.bbr.2008.08.012>.
- Han Zhang, Xi-Nian Zuo, Shuang-Ye Ma, Yu-Feng Zang, Michael P. Milham, and Chao-Zhe Zhu. Subject Order-Independent Group ICA (SOI-GICA) for Functional MRI Data Analysis. *NeuroImage*, March 2010. ISSN 10538119. doi: 10.1016/j.neuroimage.2010.03.039. URL <http://dx.doi.org/10.1016/j.neuroimage.2010.03.039>.
- Yuan Zhong, Huinan Wang, Guangming Lu, Zhiqiang Zhang, Qing Jiao, and Yijun Liu. Detecting Functional Connectivity in fMRI Using PCA and Regression Analysis. *Brain Topography*, 22(2):134–144, September 2009. ISSN 0896-0267. doi: 10.1007/s10548-009-0095-4. URL <http://dx.doi.org/10.1007/s10548-009-0095-4>.
- Yuan Zhou, Meng Liang, Lixia Tian, Kun Wang, Yihui Hao, Haihong Liu, Zhenning Liu, and Tianzi Jiang. Functional disintegration in paranoid schizophrenia using resting-state fmri. *Schizophr Res*, 97(1-3):194–205, Dec 2007. doi: 10.1016/j.schres.2007.05.029. URL <http://dx.doi.org/10.1016/j.schres.2007.05.029>.
- Yuan Zhou, Ni Shu, Yong Liu, Ming Song, Yihui Hao, Haihong Liu, Chunshui Yu, Zhenning Liu, and Tianzi Jiang. Altered resting-state functional connectivity and anatomical connectivity of hippocampus in schizophrenia. *Schizophr Res*, 100(1-3):120–132, Mar 2008. doi: 10.1016/j.schres.2007.11.039. URL <http://dx.doi.org/10.1016/j.schres.2007.11.039>.
- A. P. Zijdenbos, R. Forghani, and A. C. Evans. Automatic "pipeline" analysis of 3-D MRI data for clinical trials: application to multiple sclerosis. *IEEE Transactions on Medical Imaging*, 21(10):1280–1291, October 2002. ISSN 0278-0062.

doi: 10.1109/TMI.2002.806283. URL <http://dx.doi.org/10.1109/TMI.2002.806283>.

Xi-Nian Zuo, Ting Xu, Lili Jiang, Zhi Yang, Xiao-Yan Cao, Yong He, Yu-Feng Zang, F. Xavier Castellanos, and Michael P. Milham. Toward reliable characterization of functional homogeneity in the human brain: Preprocessing, scan duration, imaging resolution and computational space. *NeuroImage*, October 2012. ISSN 10538119. doi: 10.1016/j.neuroimage.2012.10.017. URL <http://dx.doi.org/10.1016/j.neuroimage.2012.10.017>.