

1

Université de Montréal

2

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

3

4

par

5

Christian Dansereau

6

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

7

8

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

9

10

11

Décembre, 2014

12

© Christian Dansereau, 2014.

13

Université de Montréal
Faculté des études supérieures

14

15

Cet examen pré-doctoral intitulé:

16

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

17

18

19

présenté par:

20

21

Christian Dansereau

22

23

24

a été évalué par un jury composé des personnes suivantes:

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

25

Examen accepté le:

ABSTRACT

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

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

CONTENTS

43	ABSTRACT	iii
44	CONTENTS	iv
45	LIST OF TABLES	vi
46	LIST OF FIGURES	vii
47	CHAPTER 1: GENERAL CONTEXT	1
48	1.1 Alzheimer's disease	2
49	1.2 fMRI introduction	5
50	1.3 Preprocessing	6
51	1.4 Multi-site	7
52	1.5 Connectivity	9
53	1.6 Prediction	15
54	1.6.1 Prediction in the context of AD	15
55	1.6.2 Importance of preprocessing to improve classification	15
56	1.6.3 High dimensionality problem	15
57	1.6.4 Feature selection	16
58	1.6.5 Classification	17
59	1.7 Objectives	17
60	CHAPTER 2: SCRUBBING AND MOTION ARTEFACT CORRECTION	18
61	2.1 Introduction	18
62	2.2 Method	19
63	2.2.1 Dataset	19
64	2.2.2 Preprocessing	20
65	2.2.3 Statistical analysis	21
66	2.3 Results	21

67	2.3.1	Motion distribution in CNY, CNE, pMCI and pDAT	21
68	2.3.2	Default mode network in young adult and elderly population . .	23
69	2.3.3	Impact of scrubbing on the connectivity	25
70	2.3.4	Impact of scrubbing on our discriminative power between pop- ulations	31
72	2.4	Conclusion	33
73	CHAPTER 3: MULTI-SITE CORRECTION AND FEASIBILITY	34
74	3.1	Introduction	34
75	3.2	Method	34
76	3.2.1	Preprocessing	34
77	3.2.2	Feature selection	35
78	3.2.3	Simulations	39
79	3.3	Results	41
80	3.4	Conclusion	50
81	CHAPTER 4: FUTURE DEVELOPMENT	51
82	4.1	Large feature space	51
83	4.2	Feature selection and classification	52
84	4.3	Industry application and translational effort	53
85	4.4	Timeline	54
86	REFERENCES	56

LIST OF TABLES

88	2.I	Retention table after scrubbing	22
89	3.I	Regions selected in the literature review, the region number corre-	
90		spond to the number in the Cambridge 100 partition.	45

LIST OF FIGURES

92	1.1	Biomarker model of Alzheimer's disease	4
93	1.2	Schematic of the BOLD effect	6
94	1.3	Resting-state correlation maps	11
95	1.4	Functional connectome	12
96	1.5	Consistent resting-state network (CRSN)	14
97	2.1	FD distribution among groups	23
98	2.2	Preprocessing impact on DMN for CNY and CNE	24
99	2.3	Impact of the preprocessing on the DMN (CNE,pMCI,pDAT) . .	26
100	2.4	Impact of the preprocessing on the DMN (CNY)	27
101	2.5	Impact of preprocessing on each group	28
102	2.6	Scrubbing impact on group differences	30
103	2.7	Detection power with Scrubbing	32
104	2.8	Detection power with CompCor and GSC	33
105	3.1	ICC scores	41
106	3.2	Cited region frequency	42
107	3.3	Selected region inside DMN	43
108	3.4	Selected region outside DMN	44
109	3.5	inter vs. intra site variability	46
110	3.6	DMN variability across sites	47
111	3.7	Detection power	48
112	3.8	Detection power 2	49

CHAPTER 1

GENERAL CONTEXT

115 The number of Canadians suffering from Alzheimer's disease (AD) is rapidly in-
116 creasing, with tremendous social and economic impact. Despite the emergence of promis-
117 ing drugs, the recent clinical trials with demented patients have failed. Dementia how-
118 ever comes very late in the development of the disease, at a stage where the degenera-
119 tion of neural tissues has likely gone beyond repair. In order to be efficient, therapies
120 should be initiated in the decades predating dementia, in a preclinical stage where pa-
121 tients experience no or very mild symptoms (see chapter 1.1). There are unfortunately
122 no biomarker(s) that are currently predictive of AD in this preclinical stage, and could
123 help identify the individuals that could benefit from such interventions. A promising
124 technique is resting-state functional magnetic resonance imaging (rs-fMRI), which may
125 be able to capture the early synaptic dysfunction seen in AD (see chapter 1.2). In order to
126 be able to apply statistical analysis and machine learning methods we need to preprocess
127 the data to remove as much as possible the effect of various artefacts (hardware and phys-
128 iological). The preprocessing reduces the variability of the data and therefore provides
129 more relevant and discriminative features (see chapter 1.3). In order to further improve
130 the signal to noise ratio (SNR), multiple academic groups are collaborating to pool their
131 dataset to increase the sample size of the study. Unfortunately the gain in sample size
132 comes with a new source of variability introduced by the multicentric acquisition. Site-
133 specific MRI set-ups may bias the fMRI measures, and I am thus developing procedures
134 for inter-site normalization. Account for these sources of variance are important since
135 they may bias the predictive potential of rs-functional connectivity in a multi-site setup
136 (see chapter 1.4).

137 With the objective of extracting meaningful information that could be used as biomark-
138 ers we are using connectivity measures usually based on the correlation of spontaneous
139 fluctuations in neurovascular activity from pairs of brain regions, which have been shown
140 to be sensitive to the development of the disease (see chapter 1.5). However, with about

141 10^4 recording sites in the gray matter and close to no knowledge on the early functional
142 dysfunction in AD, there is an overwhelming number of 10^8 possible connections to
143 examine as a potential diagnostic we will therefore focus on the feature selection proce-
144 dure and the stability of those selections in order to have consistent and sparse predictive
145 features discriminative of the disease(see chapter 1.6).

146 **1.1 Alzheimer's disease**

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

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

165 The use of biomarker for the early diagnostic of pathologies has a long history, with
166 many studies showing the feasibility of using AD biomarker to predict conversion from
167 MCI to AD. These studies show that individuals on the course of developing AD can be
168 identified earlier in the course of the disease by using the MCI stage with the addition

169 of imaging and cerebrospinal fluid (CSF) biomarkers to enhance diagnostic specificity
170 (Chetelat et al. 2003, Jack et al. 1999, Mattsson et al. 2009, Yuan et al. 2009). It could
171 be possible to diagnose AD after the exclusion of other forms of dementia, although a
172 formal diagnostic can only be made after a post-mortem evaluation of the brain tissue
173 (McKhann et al. 1984). This is one of the reasons why MRI based analysis and diagnos-
174 tic tools are a current research hot topic in the neurosciences community.

175 The currently dominant hypothesis in the field for the chain of events in AD patho-
176 physiology is the β -amyloid ($A\beta$)-cascade hypothesis. The proposed premise suggests
177 that interstitial $A\beta$ proteins exert a toxic effect on surrounding neurons and synapses,
178 thereby disturbing their function (Hardy and Selkoe 2002, Shankar et al. 2008). More-
179 over, recent research study suggests that, prior to neuronal death resulting in brain atro-
180 phy, disruption of functional connectivity may arise in response to a systemic problem
181 and represent an early outcome of $A\beta$ proteins in AD (Sheline and Raichle 2013). Al-
182 ready in the stage preceding aggregation of $A\beta$ fragments into amyloid plaques, there if a
183 dysfunction of synaptic transmission in many brain regions due to dimers and monomers
184 from the $A\beta$ cascade (D'Amelio and Rossini 2012). As illustrated in Sperling et al.
185 (2011) the dynamic range of fMRI (orange curve) precedes the structural changes as
186 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

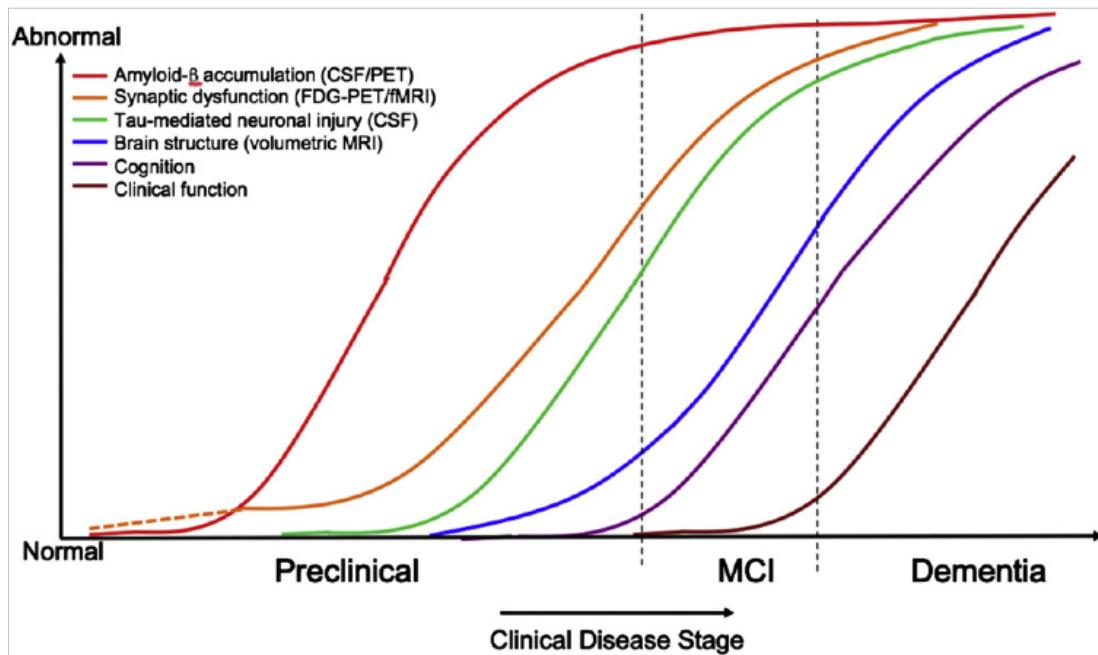


Figure 1.1: Hypothetical model of dynamic biomarkers of the AD expanded to explicate the preclinical phase: A β as identified by cerebrospinal fluid A β 42 assay or PET amyloid imaging. Synaptic dysfunction evidenced by fluorodeoxyglucose (F18) positron emission tomography (FDG-PET) or functional magnetic 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).

187 **1.2 fMRI introduction**

188 In fMRI, the acquisition process is slightly different than for the anatomical MRI
189 acquisition. fMRI uses the principle of the relaxation of hydrogen nuclei, by using spe-
190 cific fMRI sequences consisting in $T2^*$ weighted acquisitions that are sensitive to local
191 distortions of the magnetic field. Particularly, deoxyhemoglobin (a form of hemoglobin
192 without oxygen) will create such local distortions of the magnetic field, since it is a
193 paramagnetic molecule (positive magnetic susceptibility) (Ogawa et al. 1990). The
194 data acquired using fMRI rely on the hypothesis that areas showing decreased deoxy-
195 hemoglobin concentration are due to sustained brain activity. Following neuronal ac-
196 tivity, neurons require energy to restore the electrical and ionic concentration balance
197 across the cell membrane. One mechanism to generate this energy is the glucose ox-
198 idative metabolism, which requires the delivery of oxygen and glucose by the blood
199 on the site where brain activity takes place (Ogawa et al. 1990) . Initially, fMRI was
200 thought to be a good technique to measure the cerebral metabolic rate of oxygen, since
201 the new blood rushing in causes a proportional effect on the venous-end, resulting in a
202 decrease in deoxyhemoglobin concentration, and thus an increase in fMRI signal. The
203 concentration of deoxyhemoglobin actually depends mainly on three factors or phenom-
204 ena: metabolic rate of oxygen consumption, cerebral blood volume (CBV) and cerebral
205 blood flow (CBF) (Hoge et al. 1999). As a result, the fMRI signal is the outcome of
206 competing effects following neuronal activity. CBF augmentation tends to decrease the
207 deoxyhemoglobin concentration due to an increase in the oxygenated blood (resulting
208 in a increase fMRI signal). While an increase in cerebral metabolic rate of oxygen and
209 CBV tends to increase deoxyhemoglobin within the tissue at the vicinity of the neu-
210 ral activity, resulting in a decrease fMRI signal due to the paramagnetic properties of
211 deoxyhemoglobin. The fMRI signal is the result of competing effects and fortunately
212 these effects do not cancel out, allowing the detection of a net signal increase. We call
213 the interrelation of theses effects the blood oxygenation level dependent (BOLD) effect.

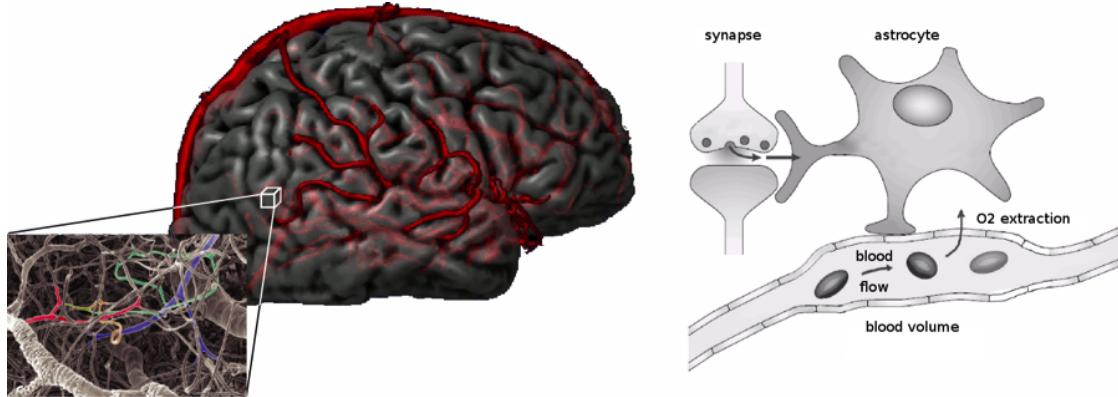


Figure 1.2: Schematic of the brain and the effect of neuronal activity on local changes in blood oxygenation signal (BOLD) (Heeger and Ress 2002)

214 1.3 Preprocessing

215 Normalization of the data is crucial to obtain a consistent and accurate classifier
 216 (Kotsiantis 2007). Therefore a particular attention is place on the correction and normal-
 217 ization procedure applied to the rs-fMRI data used in this study. A series of standard
 218 preprocessing steps is usually applied in an attempt to correct for various artefacts that
 219 would perturb the subsequent analysis. The BOLD effect associated with neuronal ac-
 220 tivity generally results in a relatively small fluctuation of the MR signal. Many factors
 221 can influence this signal. Among them, the physiological activity associated mainly
 222 with respiration, cardiac pulsations and patient's motion are major contributors to the
 223 noise and are spatially spread everywhere within the brain volume. These sources of
 224 noise result in large correlations between BOLD signals of distant voxels. Therefore,
 225 analyzing directly raw rs-fMRI data with seed-based connectivity or with a data driven
 226 technique (like ICA or clustering) will more likely result in the identification of noise
 227 related structures spread everywhere within the brain. Therefore, preprocessing meth-
 228 ods were designed in an attempt to remove specifically the so-called physiological noise
 229 and motion artefacts from the raw fMRI data. The basic steps are as follow: (1) cor-
 230 rection for slice timing differences due to delay in acquisition sampling, (2) rigid-body

231 motion estimation for within and between runs, motion correction operates by selecting
232 one functional volume as a reference to align all other functional volumes. Most head
233 motion algorithms describe head movement by 6 parameters, three translation (displace-
234 ment) parameters and rotation parameters and are appropriate to characterize motion of
235 rigid bodies. (3) Coregistration of the functional data in a reference space, (4) resampling
236 of the functional data in the stereotaxic space (references brain used as a common space
237 between subjects), (5) regression of confounds in order to remove spatially structured
238 noise on the fMRI time-series. The confounds are the slow time drift, high frequency
239 noise signal, motion parameters, the average signal white matter as well as the average
240 signal of the ventricles (containing cerebrospinal fluid CSF a frequent source of noise
241 and artefact). Some groups have suggested that these corrections are not sufficient to re-
242 move motion artefact and propose some corrective procedure (detailed in Chapter 2). (6)
243 Spatial smoothing is usually applied using a Gaussian blurring kernel to improve signal
244 to noise ratio (SNR), improved validity of the statistical tests by making the error dis-
245 tribution more normal and finally reduce anatomical and functional variations between
246 subjects (Mikl et al. 2008, Worsley and Friston 1995). There is however a few drawbacks
247 from this procedure like the reduction of spatial resolution of the data, edge artefact from
248 smoothed brain voxels with non-brain voxels which might result in hypoactivation of the
249 regions in question.

250 **1.4 Multi-site**

251 In most experiments conducted in neuroimaging, the main factors that influence
252 power are: (1) the size of the effect, determined by the difference of the means of the
253 experimental and control conditions and the variability of this difference across subjects;
254 (2) the probability of rejecting the null hypothesis when it is true; and (3) the sample size,
255 i.e. the number of subjects in the study (Desmond and Glover 2002). This last factor
256 is usually the only one controlled by the investigator this is why an increase number of
257 researchers combine their efforts in rs-fMRI to make use of multicentric data in order
258 to create large dataset suitable to statistical analysis. In research it is very difficult to

259 obtain a grant large enough to scan a cohort larger than 80 subjects, therefore researcher
260 and consortium initiatives have started to pool their resources together to make initiative
261 composed of publically available large cohorts of subjects like the 1000 functional con-
262 nectome (Biswal et al. 2010), ADNI (Mueller et al. 2005), among others. In clinical trial
263 the justification for multicentric acquisition is more of a logistical one then a financial
264 reason; they need to recruit a large amount of subject in a short period of time. In order
265 to achieve this goal they mandate the recruitment to multiple clinical centers across the
266 globe which accelerate the evaluation time of a drug. Although they may be similar by
267 their scanner protocols, scanners will have difference in their tuning and brand (even
268 field strength may differ in some cases). Unfortunately Between studies, MR acquisition
269 methodologies are among the most commonly cited sources of measurement variation
270 (Friedman et al. 2006). This is why it is important to assess if multisite resting state con-
271 nectivity analysis are feasible and what corrective measure on the data should be applied
272 to reduce the bias introduce by multisite analysis. Among the variability factor we can
273 list the following 3 categories described in (Yan et al. 2013b):

274 *Acquisition-related variations: Scanner make and model (Friedman et al. 2006),*
275 *sequence type (spiral vs. echo planar; single-echo vs. multi-echo) (Klarhofer et al.*
276 *2002), parallel vs. conventional acquisition (Feinberg et al. 2010) (Lin et al. 2005), coil*
277 *type (surface vs. volume, number of channels, orientation), repetition time, number of*
278 *repetitions, flip angle, echo time, and acquisition volume (field of view, voxel size, slice*
279 *thickness/gaps, slice prescription) (Friedman and Glover 2006)* *Experimental-related*
280 *variations: Participant instructions (Hartstra et al. 2011), eyes-open/eyes-closed (Yan*
281 *et al. 2009) (Yang et al. 2007), visual displays, experiment duration (Fang et al. 2007)*
282 *(Van Dijk et al. 2010)* *Environment-related variations: Sound attenuation measures (Cho*
283 *et al. 1998) (Elliott et al. 1999), attempts to improve participant comfort during scans*
284 *(e.g., music, videos) (Cullen et al. 2009), head-motion restraint techniques (e.g., vacuum*
285 *pad, foam pad, bite-bar, plaster cast head holder) (Edward et al. 2000) (Menon et al.*
286 *1997), room temperature and moisture (Vanhoutte et al. 2006).*

287 In 2009, the publicly released 1000 Functional Connectomes Project (FCP) and Inter-
288 national Neuroimaging Data-sharing Initiative (INDI) provided a glimpse of the variabil-

289 ity in imaging methodologies employed by the neuroimaging field. The dataset includes
290 rs-fMRI samples independently collected at imaging sites around the world, noteworthy
291 aspect of this dataset is the variation in almost every parameter of the imaging acqui-
292 sition methodologies while the majority of subject-related variables are not reported (due
293 to the majority of cases, to the fact that they were not thoroughly recorded). Despite justi-
294 fiable scepticism, feasibility analyses demonstrated that meaningful explorations of the
295 aggregate dataset, composed of 24 imaging sites for a grand total of 1093 subjects, could
296 be performed (Biswal et al. 2010) although no explicit correction for multisite variability
297 was used they only use global signal correction (GSC) to normalize subjects which may
298 introduce anti-correlation in the data (Carbonell et al. 2014, Fox et al. 2009, Murphy
299 et al. 2009, Power et al. 2014, Saad et al. 2012). After accounting for site-related differ-
300 ences, the analysis showed brain–behaviour relationships with phenotypic variables
301 such as age, gender, and diagnostic label, and confirmed a variety of prior hypotheses
302 (Biswal et al. 2010, Fair et al. 2012, Tomasi and Volkow 2010, Zuo et al. 2012). While
303 encouraging, many uncontrolled and unknown factors in the 1000 FCP remain a source
304 of concern, as they spread beyond simple site effects and can limit the datasets utility as
305 highlighted by Yan et al. (2013a). An other compelling proof of multisite bias is the
306 study reported by Nielsen et al. (2013) where they did an analysis on a single site dataset
307 and a multisite dataset of subject with autism and concluded that the multisite autism
308 study classification accuracy significantly outperformed chance but was much lower for
309 multisite prediction than for previous single site results (Nielsen et al. 2013). We there-
310 fore need to keep in mind that the site effect must be taken in account in the analysis or
311 we may reduce our detection power.

312 1.5 Connectivity

313 Resting-state (RS) functional connectivity consists in studying slow fluctuations of
314 hemodynamic activity without performing any specific task. These temporal fluctuations
315 can be monitored using the BOLD signal measured with fMRI. The first study that in-
316 troduced the concept of resting state functional connectivity was the one of Biswal et al.

317 (1995). By performing a task activation response to bilateral left and right finger tapping,
318 Biswal and colleagues were able to activate the corresponding motor areas on the cortex.
319 In a second analysis, they considered one of the area activated during the task as a seed
320 region for a seed-based analysis in resting-state condition. Seed-based analysis consists
321 in detecting temporal correlation between the signal of the predefined seed area and the
322 time course of all the voxels of the brain. Using this seed-based analysis, they found RS
323 correlations between similar brain regions than the ones involved during the task (a more
324 recent review done by Fox and Raichle (2007) illustrated in Figure 1.3 show the ability
325 to identify the complete sensorimotor network using only the bold signal from a small
326 region in that network).

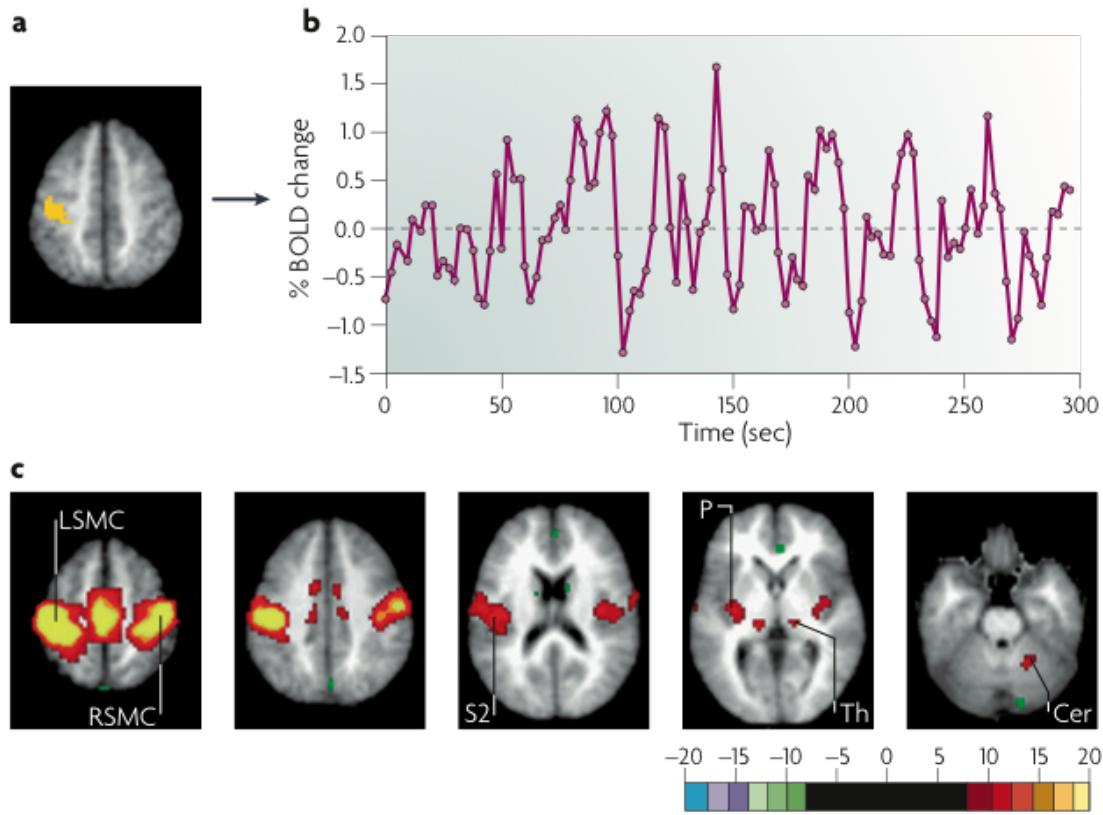


Figure 1.3: Generation of resting-state correlation maps. a) Seed region in the left somatomotor cortex (LSMC) is shown in yellow. b) Time course of spontaneous blood oxygen level dependent (BOLD) activity recorded during resting fixation and extracted from the seed region. c) Statistical z-score map 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)

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

330 al. suggest that it is possible to identify the functional organization of different struc-
 331 tures without doing any specific task. Several studies have demonstrated that patterns
 332 extracted from temporal correlations of rs-BOLD signals within the brain volume are or-
 333 ganized in space and highly reproducible from subject to subject: the so-called consistent
 334 resting state networks (CRSN) (Damoiseaux et al. 2006). Each network is a combination
 335 of multiple brain regions or units, not necessarily spatially close to each other, which are
 336 sharing similar low frequency fluctuations of the BOLD signal; this is usually repre-
 337 sented as a functional connectivity matrix where one column of the matrix represent the
 338 connectivity of a region with the rest of the brain called a functional connectivity map
 339 (see Figure 1.4 for a graphical representation of the two concepts).

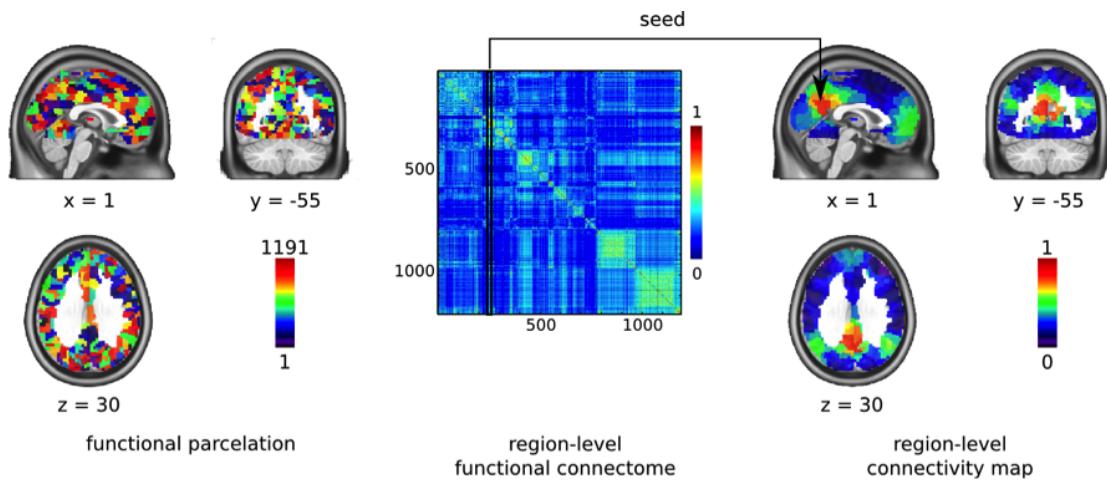


Figure 1.4: Functional connectivity; on left a representation of a functional parcellation, in the middle a region-level functional connectome representing the connectivity between each pair of region and on the right the connectivity map based on a region of interest extracted from the functional connectome

340 Several techniques have been used to identify these so-called resting-state networks
 341 (RSN). These networks show the functional organisation of various brain regions and
 342 several studies have demonstrated the relation of those functional networks to specific
 343 tasks in humans as well as in animal models (Biswal et al. 1995) (Buckner et al. 2008)
 344 (Greicius et al. 2009). Cordes et al. describe that several functional networks (sensori-

345 motor, language and visual) identified in RS condition were exhibiting similar regions
346 than the ones involved when performing a specific task (motor, language, or visual task)
347 (Cordes et al. 2000). Based on such observations, CRSNs were then labeled according
348 to the corresponding brain areas found active when performing such a specific task (e.g.
349 motor or cognitive for example) thus associated with a brain function: e.g. auditory,
350 visual, language networks (Cordes et al. 2000) see Figure 1.5 for a list of common RSN.
351 It enables us to establish the relationship between regions active when performing a spe-
352 cific task and RSNs, thus supporting the functional relevance of these networks. Subse-
353 quently, rs-fMRI signal have been used in healthy subjects to investigate normal brain
354 function, within various functional systems, such as auditory (Cordes et al. 2001), vi-
355 sual (Lowe et al. 1998), language (Hampson et al. 2002), limbic systems (Greicius et al.
356 2003, Tian et al. 2007, Wink and Roerdink 2006) and motor (Jiang et al. 2004, Lowe
357 et al. 1998). Interestingly, resting-state fMRI signals have also been used to characterize
358 the pathophysiological changes of some diseases, such as multiple sclerosis (Lowe et al.
359 2002), epilepsy (Waites et al. 2006), schizophrenia (Liang et al. 2006, Salvador et al.
360 2007, Zhou et al. 2007; 2008), attention deficit hyperactivity disorder (Tian et al. 2007,
361 Zang et al. 2007), blindness (Liu et al. 2007, Yu et al. 2007), major depression (Anand
362 et al. 2005, Greicius et al. 2007) and acute brainstem ischemia (Salvador et al. 2005).
363 Thus, we believe that resting-state fMRI will be an increasingly important modality for
364 exploring the functional abnormalities of patients with AD.

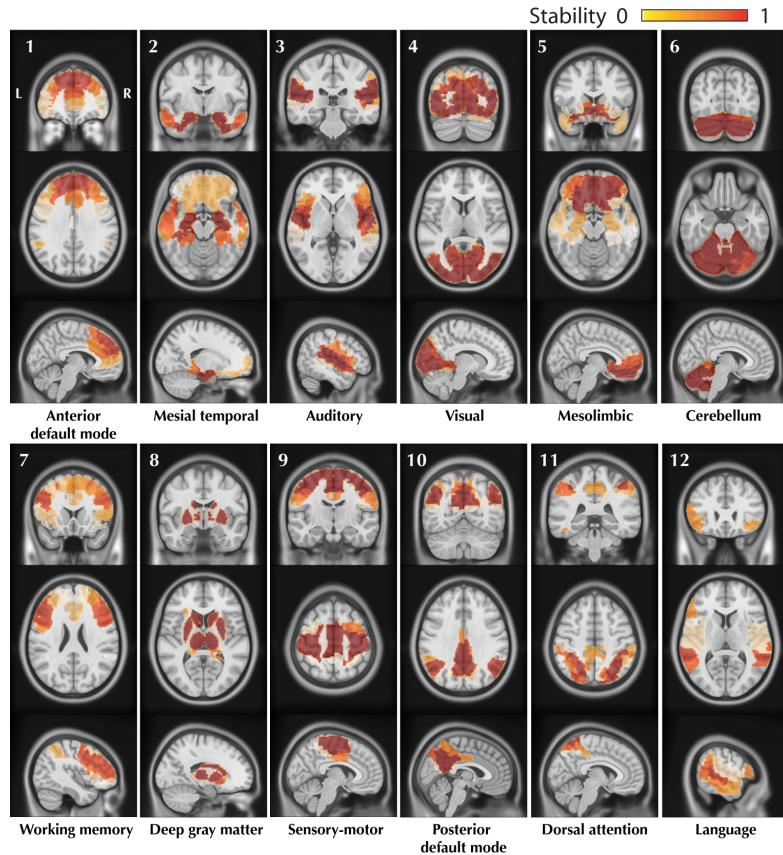


Figure 1.5: The figure shows coronal representation of 12 CRSNs identified using BASC group level analysis of 25 healthy control subjects. For each CRSN: 3 slices (coronal, axial, sagittal) are shown superimposed on an anatomical MRI template (MNI152). Labelling of each network was done visually based on previously reported CRSNs in the literature. The figure shows the usual networks: Default Mode Nettwork (#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). (?) in press

365 **1.6 Prediction**

366 **1.6.1 Prediction in the context of AD**

367 In the past few years, several major studies have been initiated that have aimed to pre-
368 dict who will develop AD, with the ultimate goal of providing a platform for therapeutic
369 intervention with disease-modifying therapies. Many of these studies were designed to
370 evaluate the role of neuroimaging and chemical biomarkers in assessing and predicting
371 progression in individuals without cognitive impairment and in individuals with MCI. A
372 very small literature currently reports findings on AD using rs-fMRI on multisite data
373 and they usually have outstanding performance like (Jiang et al. 2014) (94%) unfortu-
374 nately they did there 10 folds cross-validation after feature selection overestimating the
375 real accuracy of there analysis, the rest of the literature is mainly using leave one out
376 cross-validation and small sample for there study rising some concern with regards to
377 the generalization ability of those trained predictors (Chen et al. 2011) (87%), (Dai et al.
378 2014) (80%).

379 **1.6.2 Importance of preprocessing to improve classification**

380 Preprocessing and normalization of the source data is a crucial point that may affect
381 greatly the resulting classification or potential outcome measures. Moreover the intro-
382 duction of multicentric data in the classification will render the task even more difficult,
383 Nevertheless, a multisite cohort helps test generalizability of the results across different
384 samples, making it more likely that connections identified as discriminatory between pa-
385 tient and control reflect disease specific information rather than particularities of a single
386 dataset.

387 **1.6.3 High dimensionality problem**

388 In our modality of interest (rs-fMRI) we obtain signal from brain activity at the voxel
389 level representing more than 10^4 voxels in the gray matter cortex where the vast major-
390 ity of the neurons are located. Although preprocessed, this data continue to have a lot

391 of variance and we need to identify functional organisation more meaningful in term
392 of clinical interpretation as well as an improved representation of the feature space that
393 enhance the characterization of the functional modules. We can commonly extract func-
394 tionally and clinically meaningful network features using principal component analysis
395 (PCA) (Zhong et al. 2009), independent component analysis (ICA) (McKeown et al.
396 1998), various clustering algorithms such as k-means (Baumgartner et al. 1998), hierar-
397 chical clustering (Cordes et al. 2002), normalized cut-graph (van den Heuvel et al. 2008),
398 self-organizing maps and neural gas (Meyer-Baese et al. 2004). Once we have functional
399 parcels of the brain it is important to find the right metric to evaluate the connectivity
400 between regions of the brain at the individual level. The most common metric use is the
401 field of connectivity is the Person correlation between each pair of regions.

402 **1.6.4 Feature selection**

403 Feature selection is often a critical step prior to any learning algorithm. It reduce the
404 computational complexity of learning algorithms and expose potentially clinically mean-
405 ingful information. In many cases, this process can also improve the prediction accuracy
406 by removing redundant and irrelevant information leading to an improved signal to noise
407 ratio. Therefore feature selection is an important step in effective learning of large data
408 sets. The features selection methods are usually organized into two categories: filter
409 methods and wrapper methods (Kotsiantis 2007). The filter method evaluates the rele-
410 vance of features by looking only at the properties of the data without any knowledge on
411 the classification labels. The wrapper method assesses the goodness a feature subset us-
412 ing the performance of a learning algorithm. Another concern is the non-reproducibility
413 of reported markers (subset of features) is one of the main obstacle for the adoption of
414 such marker in a clinical setup. If the marker is indeed discriminative of the pathol-
415 ogy or of its progression we would expect the same features to be selected and exhibit
416 similar performance across various studies ("stable"). If the previous statement is not
417 true than the subset is not a good marker and may not be generalizable. This notion of
418 reproducibility is often named stability

419 **1.6.5 Classification**

420 Machine learning methods have become very popular to classify functional brain
421 images (Costafreda et al. 2009, Fu et al. 2008, Hahn et al. 2011, Marquand et al. 2008,
422 Nouretdinov et al. 2011) to discriminate them into normal or a specific pathological
423 disorder. There is several machine learning technique available to investigate classifica-
424 tion problems among the most popular we can denote support vector machine (SVM)
425 (Cortes and Vapnik 1995) which has been used in the past to categorize individual struc-
426 tural or functional brain images by differentiation of images from two groups (e.g. pa-
427 tient/control or male/female) (Fan et al. 2005, Kawasaki et al. 2007, Lao et al. 2004,
428 Mourao-Miranda et al. 2005). A second widely used method is linear discriminant anal-
429 ysis (LDA) classifier; the main advantage toward LDA is the fact that we are able to
430 natively include confounding regression in the decision model.

431 **1.7 Objectives**

432 The objectives of the theses is to 1) address as much as possible the motion confound-
433 ing effect, 2) evaluate the feasibility of multicentric fMRI analysis as well as identifying
434 normalization procedures to account as much as possible for variance introduce by this
435 type of configuration and 3) design a prediction pipeline for the data-driven identifica-
436 tion of biomarkers of AD in resting-state fMRI. The pipeline will have a feature selection
437 tool to find a highly reduced set of functional connections that are the most optimal to
438 predict the future conversion to a dementia of the Alzheimer's type in individuals with
439 mild cognitive impairment.

SCRUBBING AND MOTION ARTEFACT CORRECTION442 **2.1 Introduction**

443 Head motion is probably the most severe but also an unavoidable problem in fMRI
444 studies. The quality of the fMRI data is strongly affected by the presence of head mo-
445 tion. Has previously mention a number of groups proposed additional correction in order
446 to further remove head motion artefact. Although head motion can be corrected in the
447 image space, displacement of the head reduce the homogeneity of the magnetic field,
448 which is fine-tuned prior to functional scans for a given head position. Since head dis-
449 placements lead to non-optimal tuning, motion artefacts are not fully removed even after
450 perfect realignment of successive functional volumes in image space, we therefore need
451 to put more attention on the impact of motion and the method available to reduce its ef-
452 fect on functional connectivity (typically subjects with average motion over 1-2mm are
453 excluded from the analysis). Global signal correction (GSC); consist of regressing the
454 average of all the time-series in the brain. This technique as raise a number of concerns
455 due to the potential introduction of anti-correlation in the data (Carbonell et al. 2014, Fox
456 et al. 2009, Murphy et al. 2009, Power et al. 2014, Saad et al. 2012). Another corrective
457 measure is the CompCor method (Behzadi et al. 2007) that regresses out the n first com-
458 ponents of a PCA (principal component analysis) decomposition on the time-series of
459 the white matter and CSF-only voxels, this provided a compact representation of the time
460 series data from the noise ROI. In the two previous cases those correction are global and
461 are therefore not specific to motion correction and in some cases like GSC incorporate
462 signal from the grey matter regions. Power and colleagues showed in 2012 that spurious
463 but systematic correlations in fc-MRI arise from subject motion even at sub threshold
464 values (variation < 1 mm in motion) and are not adequately countered by common func-
465 tional connectivity processing steps. Power et al. (2012) suggested using scrubbing in
466 order to have an unbiased correction of the time-series where the frame with motion

above a certain threshold are simply removed from the time-series as well as a metric to have an index measure of the amount of head motion from one frame to the other called framewise displacement (FD). It is calculated as the sum of the absolute values of the differentiated realignment estimates at every time point giving an approximation of the motion frame by frame in millimeter. The main limitation of the scrubbing method is the fact that we reduce the number of time-points, which can lead to subjects lost due to an insufficient number of time point remaining after scrubbing correction. The question is therefore to evaluate if the trade-off of less data of greater quality is more beneficial in term of statistical power than more noisy data. I have recently published results concerning this issue Dansereau et al. (2014) (published proceeding and manuscript in preparation) in this study 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).

2.2 Method

2.2.1 Dataset

Methods The paper studies 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 (from the ADMTL dataset), 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 have 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).

492 **2.2.2 Preprocessing**

493 The datasets were analysed using the NeuroImaging Analysis Kit (NIAK¹) version
 494 0.12.14, under CentOS version 6.3 with Octave² version 3.8.1 and the Minc toolkit³ ver-
 495 sion 0.3.18. Analyses were executed in parallel on the "Mammouth" supercomputer⁴,
 496 using the pipeline system for Octave and Matlab (Bellec et al. 2010a), version 1.0.2.
 497 Brain map visualizations were created using MRICron software (Rorden et al. 2007).
 498 Each fMRI dataset was corrected of inter-slice difference in acquisition time and the pa-
 499 rameters of a rigid-body motion was estimated for each time frame. Rigid-body motion
 500 was estimated within as well as between runs, using the median volume of the first run
 501 as a target. The median volume of one selected fMRI run for each subject was coregis-
 502 tered with a T1 individual scan using Minctracc (Collins et al. 1998), which was itself
 503 non-linearly transformed to the Montreal Neurological Institute (MNI) template (Fonov
 504 et al. 2011) using the CIVET pipeline (Zijdenbos et al. 2002). The MNI symmetric tem-
 505 plate was generated from the ICBM152 sample of 152 young adults, after 40 iterations
 506 of non-linear coregistration. The rigid-body transform, fMRI-to-T1 transform and T1-
 507 to-stereotaxic transform were all combined, and the functional volumes were resampled
 508 in the MNI space at a 3 mm isotropic resolution. The 'scrubbing' method of (Power et al.
 509 2012), was used to remove the volumes with excessive motion with three cut-offs points:
 510 no scrubbing ($FD \geq 0$), $FD \geq 0.5$ and $FD \geq 0.2$. A minimum number of 50 unscrubbed
 511 volumes per run, corresponding to ~ 125 s of acquisition for a TR of 2.5 seconds, was
 512 then required for further analysis. For this reason, some subjects were rejected from the
 513 subsequent analyses: 11 CNE, 13 pMCI and 3 pDAT for a scrubbing at $FD \geq 0.5$ and
 514 83 CNE 95 pMCI and 39 pDAT for a scrubbing at $FD \geq 0.2$ (see table 2.I). The fol-
 515 lowing nuisance parameters were regressed out from the time series at each voxel: slow
 516 time drifts (basis of discrete cosines with a 0.01 Hz high-pass cut-off), average signals

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-eii>

517 in conservative masks of the white matter and the lateral ventricles as well as the first
 518 principal components (95% energy) of the six rigid-body motion parameters and their
 519 squares (Lund et al. 2006),(Giove et al. 2009). The fMRI volumes were finally spatially
 520 smoothed with a 6 mm isotropic Gaussian blurring kernel.

521 2.2.3 Statistical analysis

522 Impact of motion and scrubbing level on connectivity was estimated with a *t*-test and
 523 a group false-discovery rate (FDR) (Hu et al. 2010). Group differences (CNY, pMCI and
 524 pDAT) were investigated using *t*-tests for ten point-to-point connections selected based
 525 on a literature review (Dansereau et al. 2013) including covariates to model age, gender
 526 and site-specific bias using model averaging (Willer et al. 2010). To test the robustness of
 527 the group differences, the *t*-tests ($p \leq 0.05$) were replicated 10,000 times using random
 528 subsamples including 70% of each group.

$$\frac{\sum_{i=1}^n (pce_i \leq 0.05)}{n}, n = 10000 \quad (2.1)$$

529 Significance was assessed by a *t*-test in a linear model, including covariates to model
 530 site-specific bias.

531 2.3 Results

532 2.3.1 Motion distribution in CNY, CNE, pMCI and pDAT

533 A wide range of motion level was observed in the sample (Figure 2.1) and over all the
 534 elderly population seems to have more motion than the CNY. Even when scrubbing ag-
 535 gressively the difference in FD distribution between groups remain. We therefore need
 536 to account for FD differences in subsequent analysis. In Figure 2.1 a good overlap in
 537 the FD distribution across groups has been observed, the effect of scrubbing reduce the
 538 spread of the distribution and center the average FD around 0.2 for scrubbing with a
 539 threshold $FD \geq 0.5$ and 0.1 for a threshold of $FD \geq 0.2$. Overall the CNY subject move
 540 less compared to the elderly population as shown in the boxplot representation on the

541 left of Figure 2.1. In all preprocessing strategies the CNY remain significantly different
 542 in term of FD compared to all the elderly groups. In term of the comparissons in FD
 543 distribution among elderly poupluations, scrubbing is able to remove some of the vari-
 544 ability between groups but not suficiently to erase the differences between CNE-pMCI
 545 and CNE-pDAT and the differences between pMCI and pDAT are not significant in any
 546 preprocessing strategy. Results also show that in the elderly population, the pDAT popu-
 547 lation is the one with the smallest amount of motion falledow by pMCI and finally CNE.
 548 The is a significant difference in FD distribution between CNY and all the other groups
 549 in the 3 preprocessing strategies (see Figure 2.1 on the right). Significant differences in
 550 term of FD distribution arise when scrubbing at $FD \geq 0.5$ and $FD \geq 0.2$ between the
 551 pDAT group and pMCI as well as pDAT and CNE group.

552 In term of the impact of scrubbing on the retention of the original dataset after the
 553 procedure we obtain 90% of subjects who survived the exclusion criteria of a minimum
 554 of 50 frames after scrubbing at $FD \geq 0.5$ and 30% surviving rate with an aggressive
 555 scrubbing of $FD \geq 0.2$ see Table 2.I for the elderly poupluation (CNE, pMCI and pDAT).
 556 Only 1% of the subjects were lost in the CNY cohort (99% survival) for an $FD \geq 0.5$
 557 and 65% survived for $FD \geq 0.2$.

	CNE	pMCI	pDAT	Total	CNY
FD≤0	126	133	54	313	355
FD≤0.5	115	120	51	286	352
FD≤0.2	43	38	15	96	230
Retention :FD≤0.5	91%	90%	94%	91%	99%
Retention :FD≤0.2	34%	29%	28%	31%	65%

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

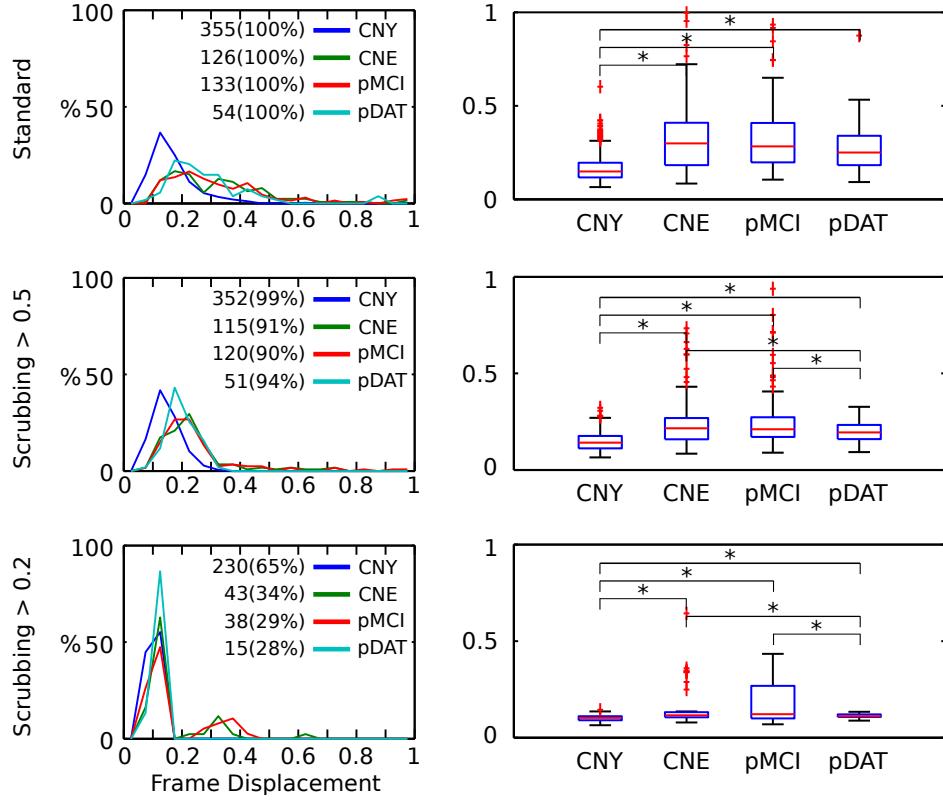


Figure 2.1: Distribution of the frame displacement (FD) for 3 groups (CNE, pMCI, pDAT) when scrubbing is applied at various levels (no scrubbing, scrubbing of $FD > 0.5$ and scrubbing of $FD > 0.2$). The boxplot on the right show the distribution of FD with there associated statistical differences t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

558 2.3.2 Default mode network in young adult and elderly population

559 A visual inspection of the DMN in standard preprocessing strategy show more pos-
 560 itive correlation values in the CNE population compared to the CNY (See Figure 2.2).
 561 Moreover we have a slight decrease in connectivity in the frontal part of the DMN in
 562 the CNE population compared to the more common pattern of fronto parietal connec-
 563 tivity shown in the CNY population. This finding of more negative correlation for CNE
 564 compared to CNY can be observed in all the preprocessing strategies except for GSC
 565 Figure 2.2, who seams to increase the extent of the negative correlation found in other

566 preprocessing strategies.

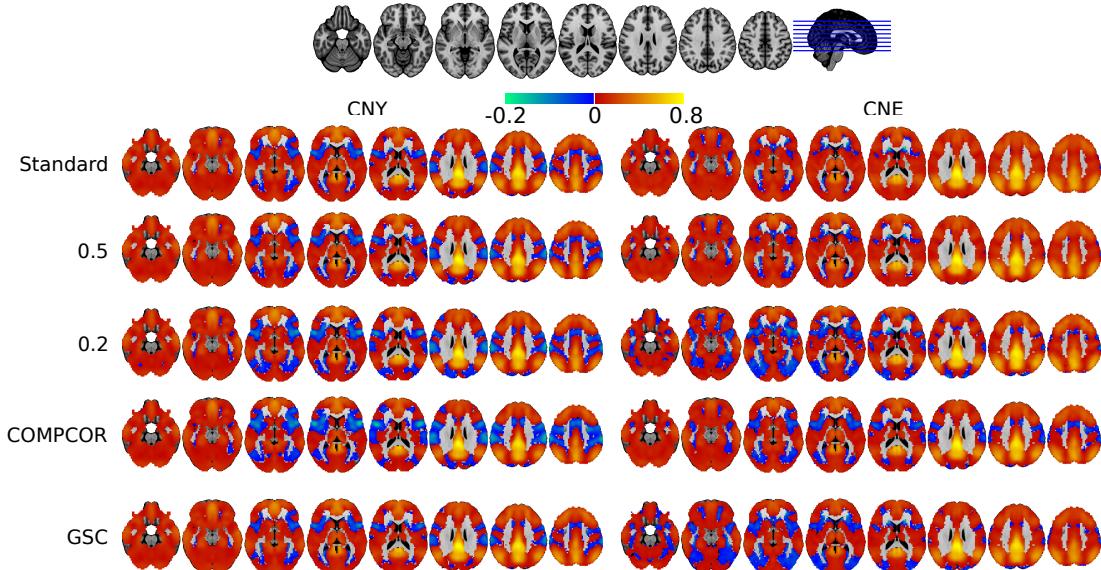


Figure 2.2: Overlay of the average default mode network (DMN) with various preprocessing strategies for CNY and CNE populations on the ICBM 152 anatomical atlas. Seed based maps with a seed in the PCC using Fisher z transform of the Pearson's r correlation between the average time series of each network.

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

576 **2.3.3 Impact of scrubbing on the connectivity**

577 **2.3.3.0.1 Global impact of scrubbing** Scrubbing significantly increased connectiv-
578 ity strength inside the default-mode network and reduced connectivity with anti corre-
579 lated regions. Difference in functional connectivity of the DMN show significant in-
580 crease of the frontal part of the DMN with scrubbing at $FD \geq 0.5$ and $FD \geq 0.2$ a region
581 normally positively correlated with the PCC (see 2.3). Significant decrease of connec-
582 tivity with region associated with attention are also observed (dorsal attention network).
583 The preprocessing with CompCor show massive decrease in the sensorimotor region
584 and globally across the brain except for a mesio frontal region (anterior cingulate cor-
585 tex ACC) more ventral then the expected mesio frontal region associated with the DMN
586 (medial prefrontal cortex MPFC). The map of difference for the GSC show massive de-
587 crease in connectivity across the brain and in particular the occipital lobe, premotor and
588 sensorimotor areas. The effect reported for the preprocessing with GSC are even more
589 pronounced for the CNY group (see supplementary material Figure 2.4).

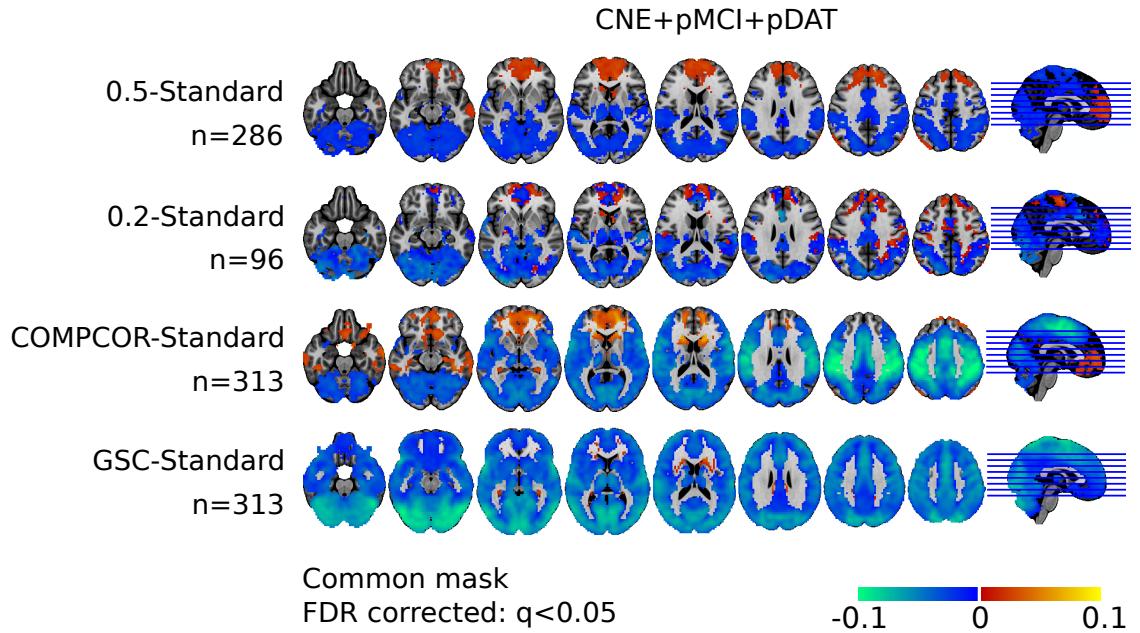


Figure 2.3: Differences in functional connectivity for the default mode network (seed in the PCC). Differences in connectivity between all the groups pooled together (CNE, pMCI, pDAT) with scrubbing ($FD > 0.5$ and $FD > 0.2$), CompCor, and GSC. The mask used depicts only significant result of the t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

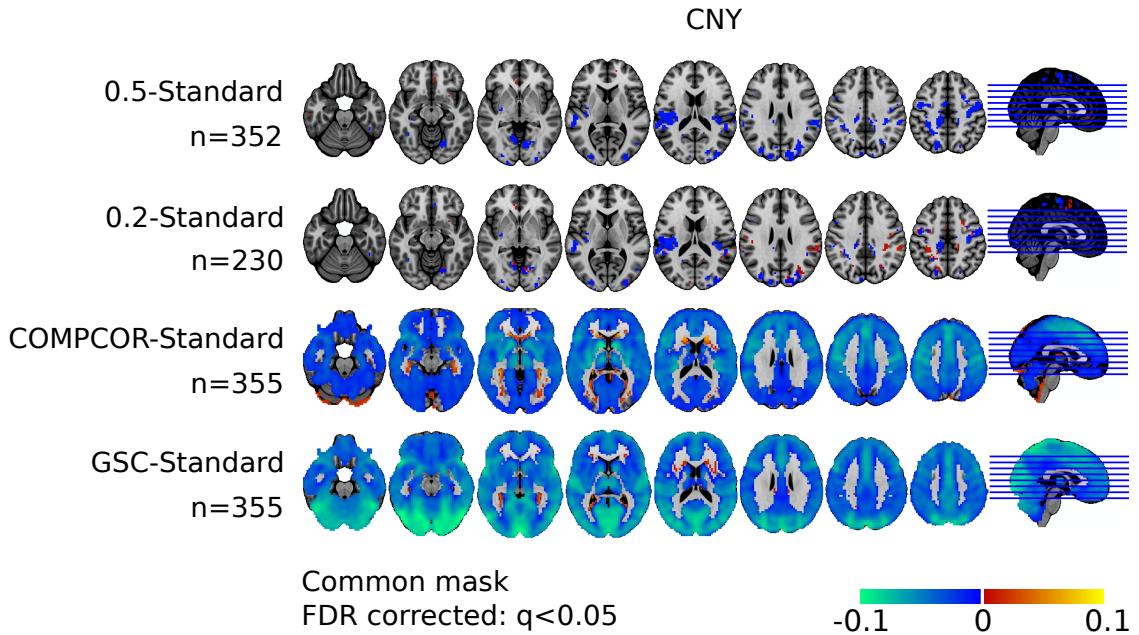


Figure 2.4: Differences in functional connectivity for the default mode network (seed in the PCC). Differences in connectivity between all the CNY with scrubbing ($FD > 0.5$ and $FD > 0.2$), CompCor, and GSC. The mask used depict only significant result of the t -test (FDR correction $q < 0.05$) only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

590 **2.3.3.0.2 Population specific impact of scrubbing** The population specific differ-
 591 ence in connectivity (see Figure 2.5) show almost identical findings as the combination
 592 of CNE, pMCI and pDAT therefore confirming that the observation are transferable in
 593 every population and not specific. Although the increase connectivity of the MPFC is
 594 observed in all groups when scrubbing is applied, the difference in connectivity is greater
 595 as we progress toward dementia (difference in fc for MPFC area CNE < pMCI < pDAT
 596). Difference map for CompCor and GSC revealed an invers patern where the greatest
 597 changes are observed in the CNE group followed by pMCI and finally pDAT for nega-
 598 tive differences. CompCor show a increase in connectivity for the ACC area in the three
 599 groups but particularly in the pMCI group followed by CNE and then pDAT.

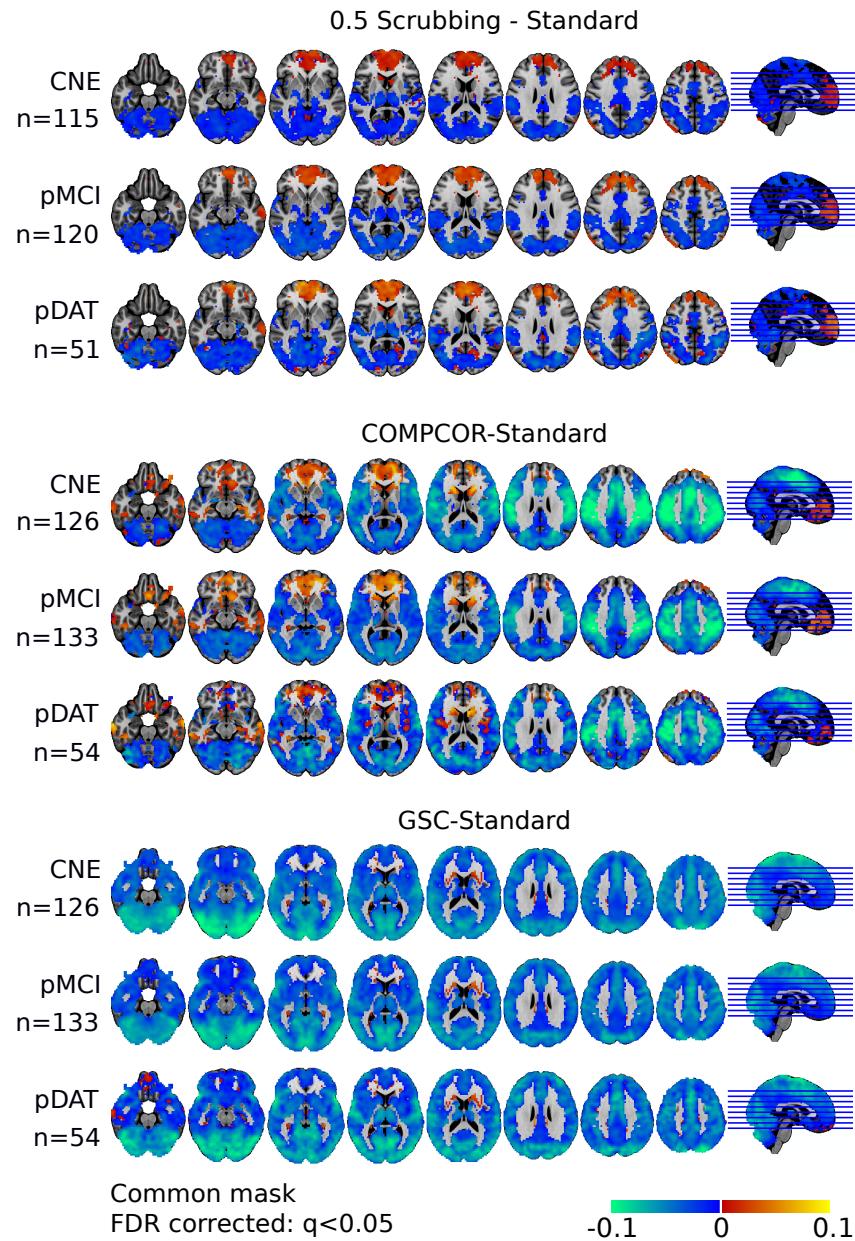


Figure 2.5: Differences in functional connectivity for the default mode network (seed in the PCC). Differences in connectivity for each group (CNE, pMCI, pDAT) compared to baseline (standard preprocessing) with and without scrubbing ($FD > 0.5$ and $FD > 0.2$) (FDR correction $q < 0.05$) for all voxels showing a significant effect in at least one of the contrast.

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

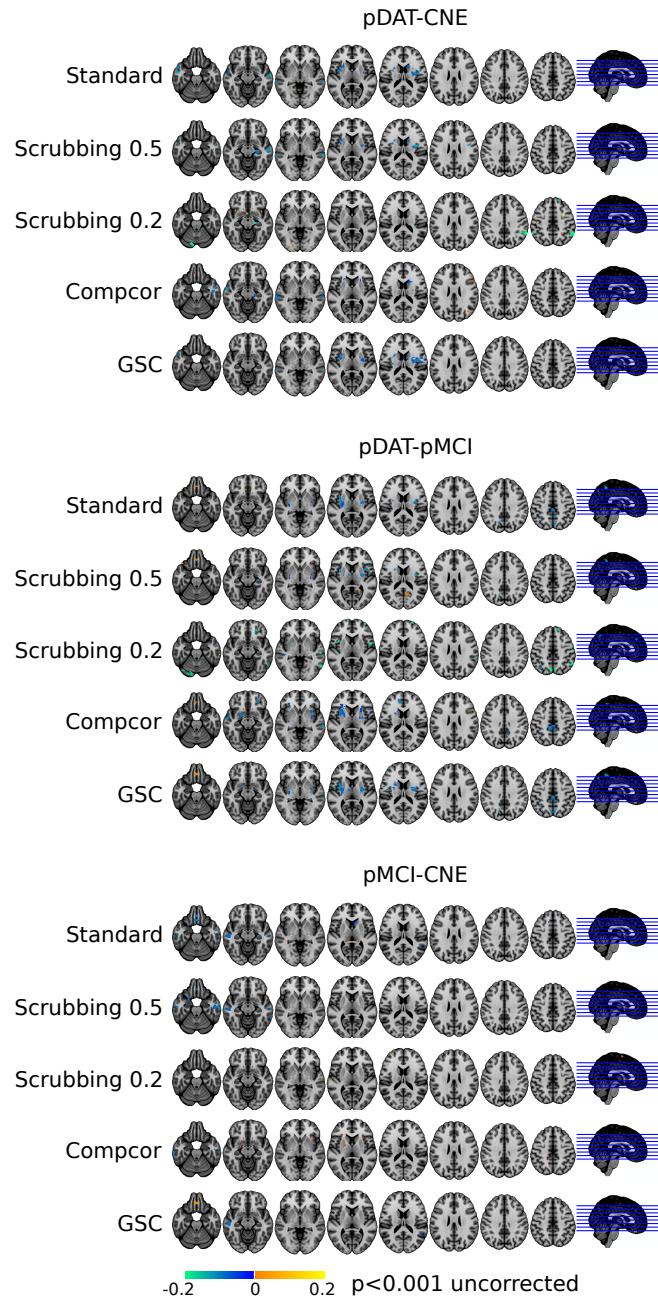


Figure 2.6: Impact of preprocessing on connectivity differences between the three groups (CNE, pMCI, pDAT). Connectivity differences between DMN (t -test $p < 0.001$ uncorrected) computed with various preprocessing strategies (standard, scrubbing (0.5 and 0.2), CompCor and GSC). The maps are represented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

603 **2.3.4 Impact of scrubbing on our discriminative power between populations**

604 **2.3.4.0.3 point to point connections** Despite a moderate loss of subjects using a
 605 scrubbing at 0.5 (due to insufficient remaining data < 50 frames), the gains in data qual-
 606 ity translated into an increased detection rate for group differences in almost all tested
 607 connections. As shown in Figure 2.7 the $FD \geq 0.5$ scrubbing procedure mitigated mo-
 608 tion artefacts and improved or maintain the statistical power of cross-sectional compari-
 609 son of elderly clinical cohort. For $FD \geq 0.2$ some improvement can be observed but the
 610 sample size remain too small to be significant in most cases. The contrast with the
 611 most improvement due to scrubbing is the pDAT-CNE which is the contrast on which
 612 the literature point to point connections was selected for. For the pDAT-CNE contrast
 613 the most improved and consistent connections are connections with the IPL and the right
 614 SFG as well as the IPL and the dMPFC3. Connection with the PCC are also markedly
 615 improved by scrubbing namely the PCC and PCUN with MTL one of the first connec-
 616 tion reported to be affected in the earliest stages of Alzheimer disease. Note that not
 617 all connection who had a good consistency with standard preprocessing improved us-
 618 ing scrubbing. Statistical power for the pMCI-CNE contrast show small improvement in
 619 PCC-PCUN, aMPFC-PCUN, IPL-dMPFC3 for $FD \geq 0.5$ and improved consistency for
 620 SFGr-FUS with a scrubbing at $FD \geq 0.2$. Finally the pDAT-pMCI contrast show gen-
 621 eral decrease statistical power when scrubbing is applied, the only exception is aMPFC-
 622 PCUN who show improvement when scrubbing $FD \geq 0.2$ is applied.

623 **2.3.4.0.4 Detection power for every scrubbing strategy** The simulation applied
 624 using CompCor and GSC preprocessing strategies revealed less consistency with the
 625 previously mention pairs of connections especially for the pDAT-CNE contrast who is
 626 supposed to be the optimised contrast for that selection of connection pairs. For the
 627 pMCI-CNE contrast GSC as consistently outperformed CompCor and Standard pre-
 628 processing in the most dominant connection namely dMPFC-dMPFC2, PCC-PCUNm
 629 and aMPFC-MTL. On the other hand the last contrast pDAT-pMCI revealed the Com-
 630 pCor preprocessing as the the preprocessing strategy with the most improved statistical
 631 power for aMPFC-PCUN, PCC-MTL and IPL-MTL. Interestingly the scrubbing proce-

632 dure seems to improve the detection power more consistently and over larger number of
 633 pair of connections than CompCor and GSC method (see Figure 2.7 and 2.8).

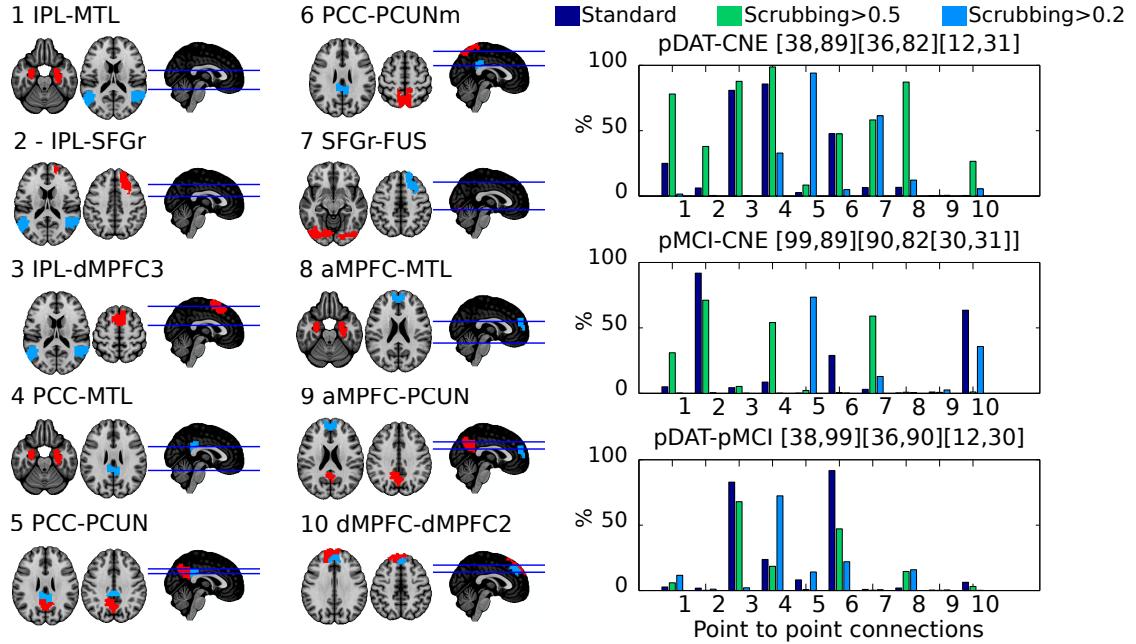


Figure 2.7: On the Left: Seeds from 10 point-to-point connections. PCC (posterior cingulate cortex), PCUN (precuneus), dMPFC (dorsomedial prefrontal cortex), dMPFC2 (dorsomedial prefrontal cortex2), IPL (inferior parietal lobule), SFGr (right superior frontal gyrus), aMPFC (anterior medial prefrontal 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 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.

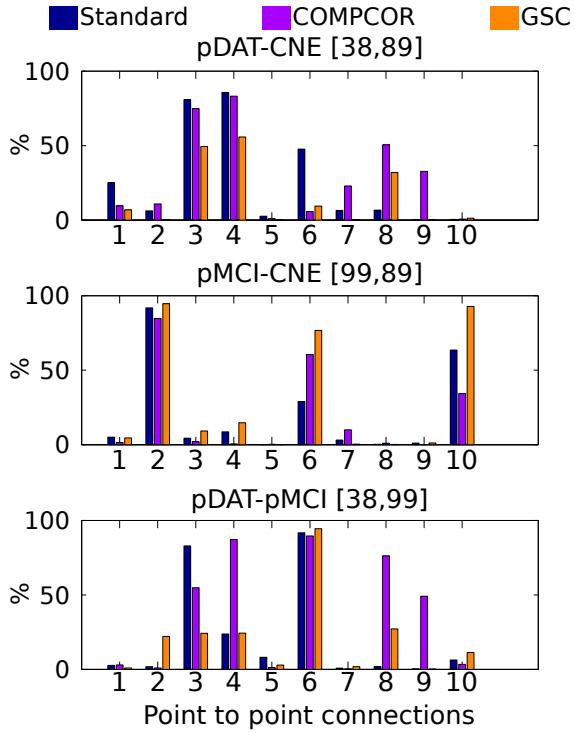


Figure 2.8: Detection power of group differences for 3 preprocessing strategy (standard, CompCor and GSC). The detection power is computed using a t -test of each connection ($p < 0.05$) replicated 10,000 times using random subsamples of 70% of each group. Explanatory variables included age and gender and a multi-site bias correction are applied.

634 2.4 Conclusion

635 Motion introduces a systematic bias in the measures of resting-state connectivity in
 636 elderly population. The scrubbing procedure mitigated motion artefacts and improved
 637 the statistical power of group-level analysis.

CHAPTER 3

MULTI-SITE CORRECTION AND FEASIBILITY

640 I have recently published results addressing the multi-site bias and the methods that
641 can be used to compensate for multi-site bias in Dansereau et al. (2013) (published pro-
642 ceeding and manuscript in preparation).

643 **3.1 Introduction**

644 Resting-state (RS) connectivity in fMRI is a promising biomarker for a variety of
645 neurological diseases. Typically in a clinical trial, a large cohort is collected from mul-
646 tiple sites. The main issue is the lack of consistency in the multi-site RS connectivity
647 acquisitions that may obscure clinically relevant information. Therefore the aims of the
648 study were to: (1) characterize the amplitude of the site bias, i.e. the systematic differ-
649 ences in rs-fMRI connectivity across different acquisition sites. (2) Quantify the impact
650 of the between-site variance on the power of statistical tests in resting-state fMRI.

651 **3.2 Method**

652 **3.2.1 Preprocessing**

653 The datasets were analysed using the NeuroImaging Analysis Kit (NIAK¹) version
654 0.12.14, under CentOS version 6.3 with Octave² version 3.8.1 and the Minc toolkit³ ver-
655 sion 0.3.18. Analyses were executed in parallel on the "Mammouth" supercomputer⁴,
656 using the pipeline system for Octave and Matlab (Bellec et al. 2010a), version 1.0.2.
657 Brain map visualizations were created using MRICron software Rorden et al. (2007).

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-eii>

658 Each fMRI dataset was corrected of inter-slice difference in acquisition time and the pa-
659 rameters of a rigid-body motion was estimated for each time frame. Rigid-body motion
660 was estimated within as well as between runs, using the median volume of the first run
661 as a target. The median volume of one selected fMRI run for each subject was coregis-
662 tered with a T1 individual scan using Minctracc (Collins et al. 1998), which was itself
663 non-linearly transformed to the Montreal Neurological Institute (MNI) template (Fonov
664 et al. 2011) using the CIVET pipeline (Zijdenbos et al. 2002). The MNI symmetric tem-
665 plate was generated from the ICBM152 sample of 152 young adults, after 40 iterations
666 of non-linear coregistration. The rigid-body transform, fMRI-to-T1 transform and T1-
667 to-stereotaxic transform were all combined, and the functional volumes were resampled
668 in the MNI space at a 3 mm isotropic resolution. The a censoring method described
669 in (Power et al. 2012) called "scrubbing" was used to remove the volumes with exces-
670 sive motion using a cut-off value of $FD \geq 0.5$. A minimum number of 50 unscrubbed
671 volumes per run, corresponding to ~ 125 s of acquisition for a TR of 2.5 seconds, was
672 then required for further analysis. The following nuisance parameters were regressed out
673 from the time series at each voxel: slow time drifts (basis of discrete cosines with a 0.01
674 Hz high-pass cut-off), average signals in conservative masks of the white matter and the
675 lateral ventricles as well as the first principal components (95% energy) of the six rigid-
676 body motion parameters and their squares (Lund et al. 2006),(Giove et al. 2009). The
677 fMRI volumes were finally spatially smoothed with a 6 mm isotropic Gaussian blurring
678 kernel.

679 3.2.2 Feature selection

680 Regions are routinely defined using an anatomical parcellation (He et al. 2009), such
681 as the AAL template (Tzourio-Mazoyer et al. 2002). Anatomical parcels may however
682 not well match the organization of resting-state networks. A framework developped
683 in the lab was used to generate data-driven functional decomposition into resting-state
684 networks based on the coherence of BOLD activity at the individual or group level (Bel-
685 lec 2006) (Bellec et al. 2010b). When a low number of networks (or scale) is used, this
686 technique, called bootstrap analysis of stable clusters (BASC), generates decompositions

687 of the brain into distributed large-scale networks, such as the DMN. At high scales, it
688 identifies subnetworks and functional regions (Kelly et al. 2012). We generated a BASC
689 parcellation in 100 clusters on the Cambridge sample, including ~200 young adults from
690 the 1000 functional connectome database (Biswal et al. 2010) and used it to generate the
691 rs-fMRI outcome measures.

692 **3.2.2.0.5 The default-mode network** Since the seminal work of Greicius et al. (2004),
693 many rs-fMRI studies in AD focused on the default-mode network (DMN), a group of
694 regions consistently more active at rest than during a broad range of different tasks (Gus-
695 nard and Raichle 2001). The DMN was notably reported to largely overlap with the
696 regions that show high amyloid-beta deposition in patients with DAT (Buckner et al.
697 2009). It includes the posterior cingulate cortex (PCC) / Precuneus (PCUN) area, the
698 inferior parietal lobule (IPL), the anterior cingulate cortex / medial prefrontal cortex
699 (MPFC) (Greicius et al. 2003). Other structures such as the medial temporal cortex or
700 the superior frontal gyrus are also generally regarded as part of different subnetworks of
701 the DMN (Andrews-Hanna et al. 2010, Margulies et al. 2009).

702 **3.2.2.0.6 Literature review: Alzheimer's disease and resting-state fMRI** We per-
703 formed a literature review to select candidate connections that have been shown to be
704 prominently impacted in Alzheimer's disease. There is no single authoritative reference
705 on the effect of a DAT on rs-fMRI connectivity, and the field has been dominated thus
706 far by studies with small samples ($n \sim 20$) and limited statistical power, see Sheline and
707 Raichle (2013) for a recent review. Because the DMN has been most extensively stud-
708 ied, we decided to focus on this network and to run a meta-analysis on six papers that
709 (1) used some analogue of seed-based connectivity maps in resting-state fMRI using one
710 or multiple seeds in the DMN (2) investigated abnormalities in resting-state functional
711 connectivity in patients suffering of a dementia of the Alzheimer's type and (3) provided
712 tables of coordinates in stereotaxic space for the results.

713 **3.2.2.0.7 review: Alzheimer's disease and resting-state fMRI**

- 714 ● Zhang et al. (2009) used functional connectivity maps with a seed in the poste-
715 rior cingulate cortex (PCC) to explore the differences between a group of elderly
716 cognitively normal subjects (CNE, n=16) and patients with a mild dementia of the
717 Alzheimer's type (DAT, n=18).
- 718 ● Zhang et al. (2010) generalized the Zhang et al. (2009) study with CNE (n=16)
719 and a larger group of patients with DAT (n=46). Patients were separated in three
720 groups (mild, moderate, severe DAT), and each group of patients was contrasted
721 against the CNE.
- 722 ● Wang et al. (2006) used functional connectivity maps with a seed in the hip-
723 pocampi to explore the differences between a group of CNE (n=13) and patients
724 with a mild DAT (n=13). All results included in the meta-analysis are from Ta-
725 ble 2, seeded in the right hippocampus. Seeds were manually delineated on an
726 individual basis.
- 727 ● Wang et al. (2007) used functional connectivity maps with a seed in the posterior
728 cingulate cortex (PCC) as well as full brain point-to-point correlations (based on
729 an AAL parcellation) to explore the differences between a group of elderly cogni-
730 tively normal subjects (CNE, n=14) and patients with a very mild to mild dementia
731 of the Alzheimer's type (DAT, n=14). Only the results based on the PCC seed were
732 included in the meta-analysis.
- 733 ● Goveas et al. (2011) used functional connectivity maps with a seed in the hip-
734 pocampi to explore the differences between a group of elderly cognitively normal
735 subjects (CNE, n=18) and patients with a mild dementia of the Alzheimer's type
736 (DAT, n=14) before and after donepezil treatment. Seeds were manually delin-
737 eated on an individual basis, before and after treatment.
- 738 ● Damoiseaux et al. (2012) used dual-regression independent component analysis
739 to explore longitudinal differences between a group of CNE (n=18) and patients
740 with DAT (n=21). All results included in the meta-analysis are from Table 3 (dif-

741 ferences at baseline) and Table 4 (interaction with time). The authors used three
742 components representing the Anterior DMN, Ventral DMN and Posterior DMN.

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

764 The TRT reliability study was based on the publicly available NYU-TRT database.
765 The database included 25 young healthy adults, and each subject had three rs-fMRI run:
766 two in a single session (separated by 45 mns) and another run 5 – 16 months latter.
767 Several outcome measures were generated in key regions impacted by AD. Using the
768 three runs, one intra-class correlation (ICC) was generated intra-session, and two ICCs
769 were generated inter-session for each outcome measures. The outcome measures were
770 ranked based on average of intra- and inter-session ICCs.

771 **3.2.3 Simulations**

772 **3.2.3.0.8 Dataset** In order to simulate various scenarios within the context of a multi-
 773 site setting, a cohort of subjects acquired at a single-site was selected to act as our ref-
 774 erence dataset and for the multi-site configuration a cohort from a collection of 7 small
 775 sites, roughly totalling the same sample size as the reference dataset, was used. The co-
 776 hort used for the study contains 385 participants from the 1000 Functional Connectomes
 777 Project (Biswal et al. 2010) (150 males, age range = 18-46 yrs) composed of 1 large site
 778 (Cambridge n=~200) and 7 small sites (n=~20/site for a total ~200). The fMRI datasets
 779 were preprocessed with the Neuro-Imaging Analysis Kit (NIAK) as described earlier in
 780 Section 3.2.1.

781 **3.2.3.0.9 Connectome** Using a brain partition of R networks obtain from BASC pro-
 782 cedure described in Bellec et al. (2010b), and taking each pair of distinct networks i and
 783 j , the between-network connectivity $y_{i,j}$ is measured by the Fisher transform of the Pear-
 784 son's correlation between the average time series of the network. The within-network
 785 connectivity $y_{i,i}$ is the Fisher transform of the average correlation between time series of
 786 every pair of distinct voxels inside network i . The connectome $\mathbf{Y} = (y_{i,j})_{i,j=1}^R$ is thus a
 787 $R \times R$ matrix. Each column j (or row, as the matrix is symmetric) codes for the connec-
 788 tivity between network j and all other brain networks (full brain functional connectivity
 789 map). For a scale with R parcels, there are exactly $L = R(R + 1)/2$ distinct elements in
 790 an individual connectome \mathbf{Y} .

791 **3.2.3.0.10 Effect size (cohen's d)** For each site and each sample, half of the subjects
 792 were randomly assigned to a "treatment" group and a Monte-Carlo simulation was used
 793 to estimate the detection power in the single-site and in multi-site setting.

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

$$s = \sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}} \quad (3.1)$$

$$d = \frac{\bar{x}_1 - \bar{x}_2}{s}$$

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

804 **3.2.3.0.11 multi-site correction approaches** For the multi-site three flavours were
 805 computed: multi-site no correction, multi-site with dummy variables and multi-site with
 806 METAL correction. Depending on the multi-site configuration and distribution of the
 807 subject we proposed two corrective approaches that can be applied as shown in the sim-
 808 ulations of Figure 3.8. The first one is the introduction of dummy variables (binary
 809 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:

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

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

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

814 The second approach is to compute the GLM independently on each site and then
 815 combine the statistical results from each site in a global score. This model averaging
 816 technique called METAL from Willer et al. (2010) model site specific bias by running a
 817 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 3.4. This is the most flexible way to account for multi-site effect while keeping the analysis simple and robust to unbalanced sites.

β_v effect size estimate for site v

se_v standard error for site v

$$\begin{aligned} w_v &= \frac{1}{se_v^2} \\ SE &= \sqrt{\frac{1}{\sum_v w_v}} \\ \beta &= \frac{\sum_v \beta_v w_v}{\sum_v w_v} \\ Z &= \frac{\beta}{SE} \\ p &= 2(1 - \phi(|Z|)) \end{aligned} \tag{3.4}$$

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).

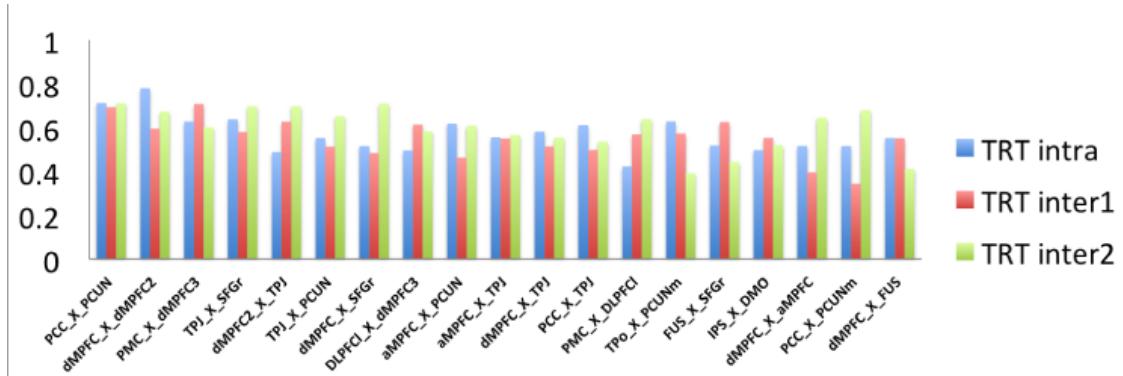


Figure 3.1: ICC scores for pairs of connections passing the > 0.5 threshold on the NYU-TRT dataset.

Average connectivity maps associated with the DMN nodes as well as the nodes

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

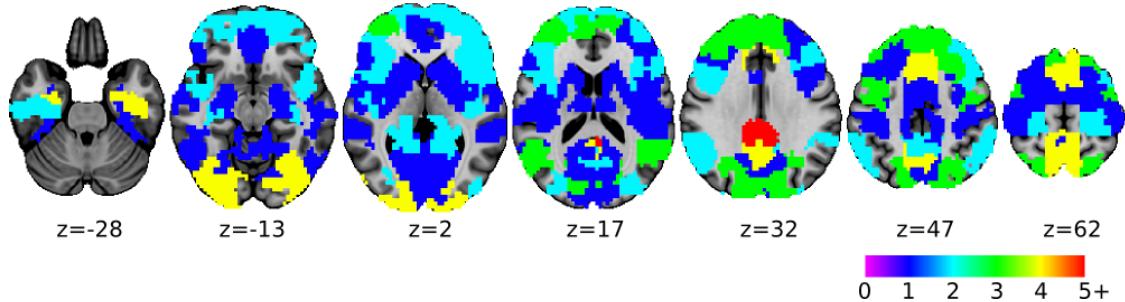


Figure 3.2: Frequency of reported regions showing functional differences based on a literature review of 6 papers.

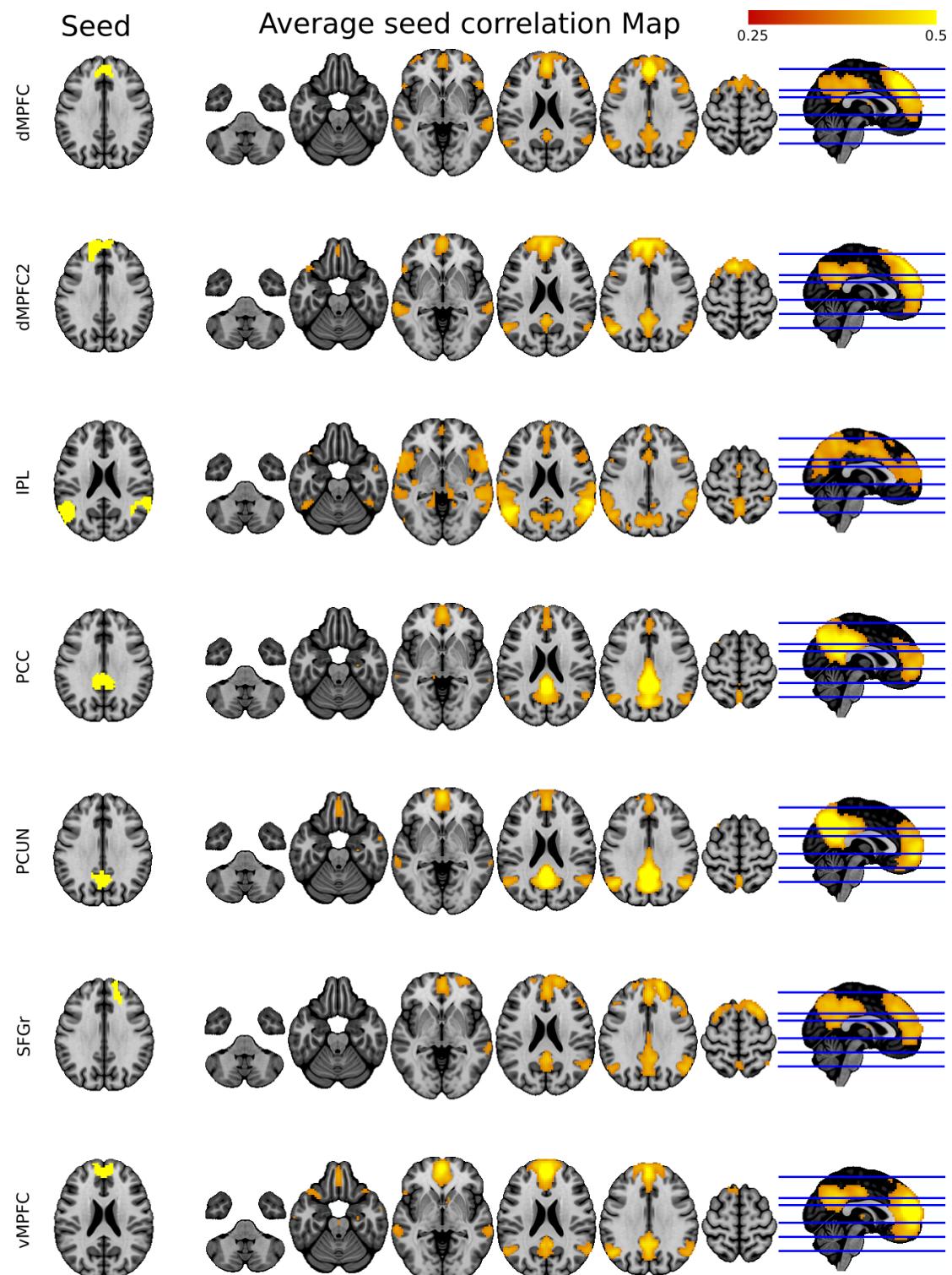


Figure 3.3: Selected nodes inside the DMN who passed the TRT selection.

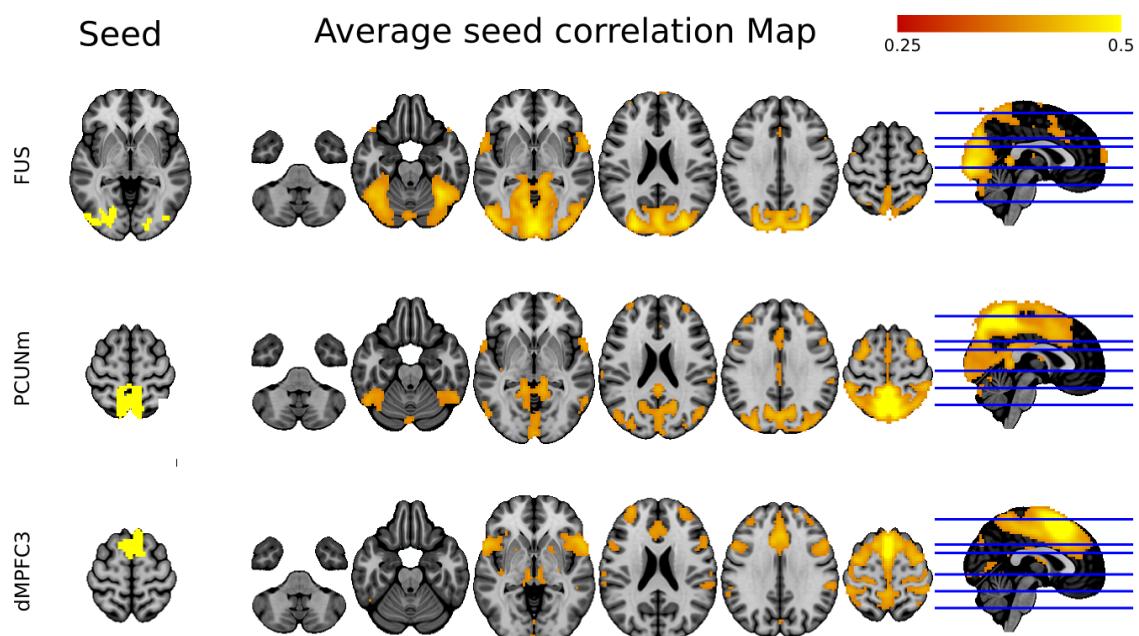


Figure 3.4: Selected nodes outside the DMN who passed the TRT selection.

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.

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

- 838 • PCC x PCUN
- 839 • dMPFC x dMPFC2
- 840 • IPL x SFGr
- 841 • aMPFC x PCUN

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

- 844 • FUS x SFGr
- 845 • PCC x PCUNm
- 846 • IPL x dMPFC3

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

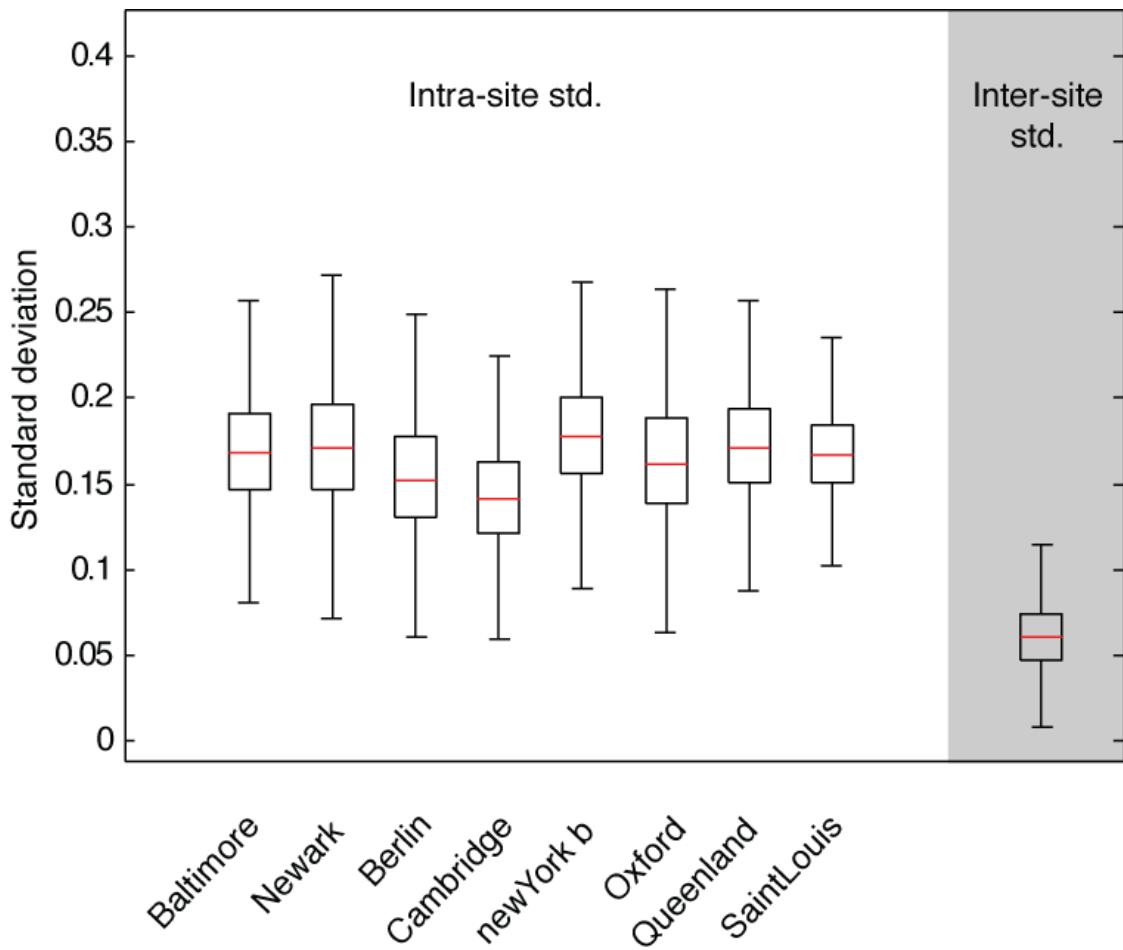


Figure 3.5: Distribution of intra-site (between-subject) standard deviation vs. inter-site (between-site) standard deviation, based on the standard deviation of the connectivity matrices from 8 sites from the 1000 functional connectome dataset.

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

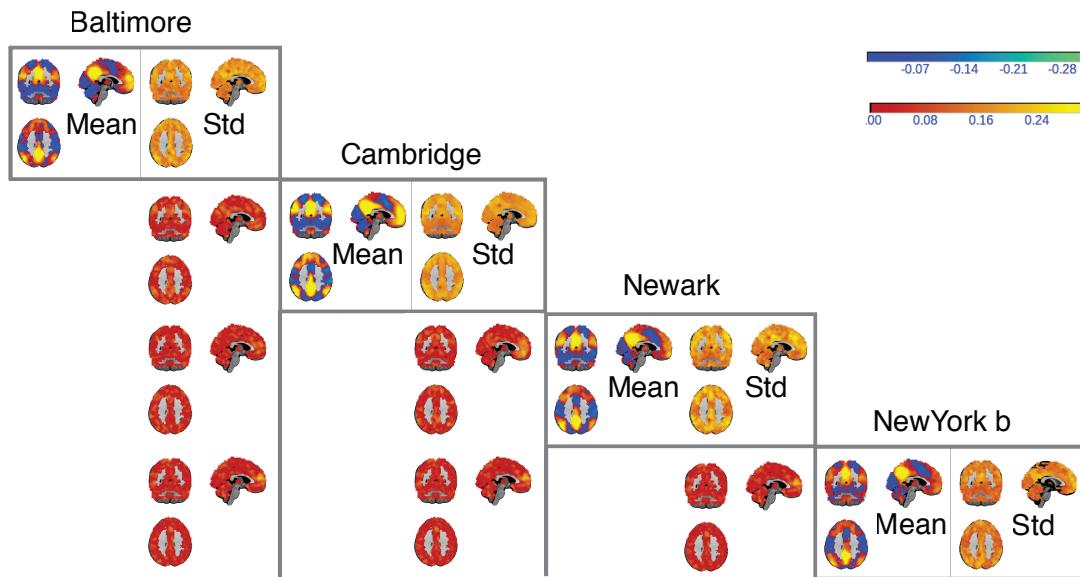


Figure 3.6: Functional connectivity maps of the default-mode network at multiple sites. The average connectivity map for 4 sites (Baltimore, Cambridge, Newark and New-Yorkb at 3T) are presented on the diagonal (left). The standard deviation across subjects and within site is presented next to it (diagonal 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).

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

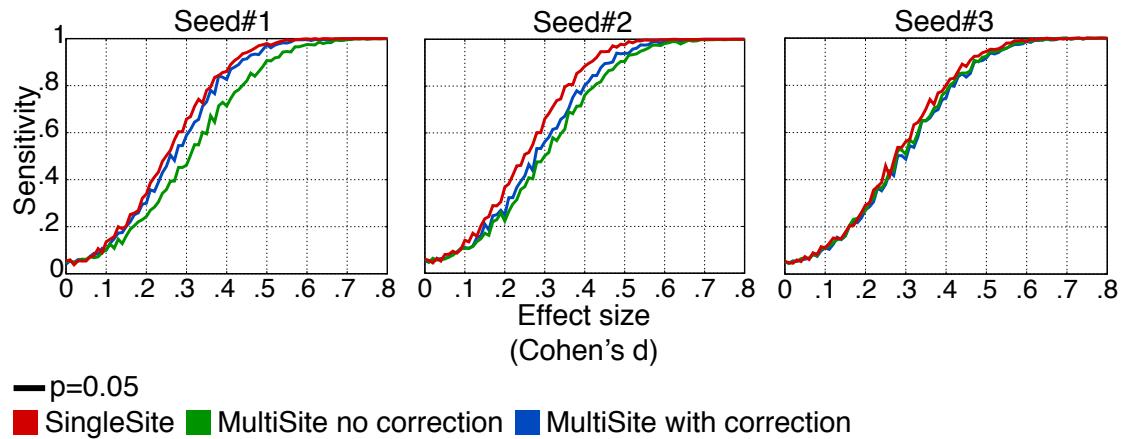


Figure 3.7: Power analysis for a resting-state fMRI study. A Monte-Carlo simulation was implemented to evaluate the power of a resting-state multi-site study, based on real values of three connections in the default-mode network PCC/MFC, rHIP/vMFC and lIIPC-rDLPFC. For each site and each sample, half 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.

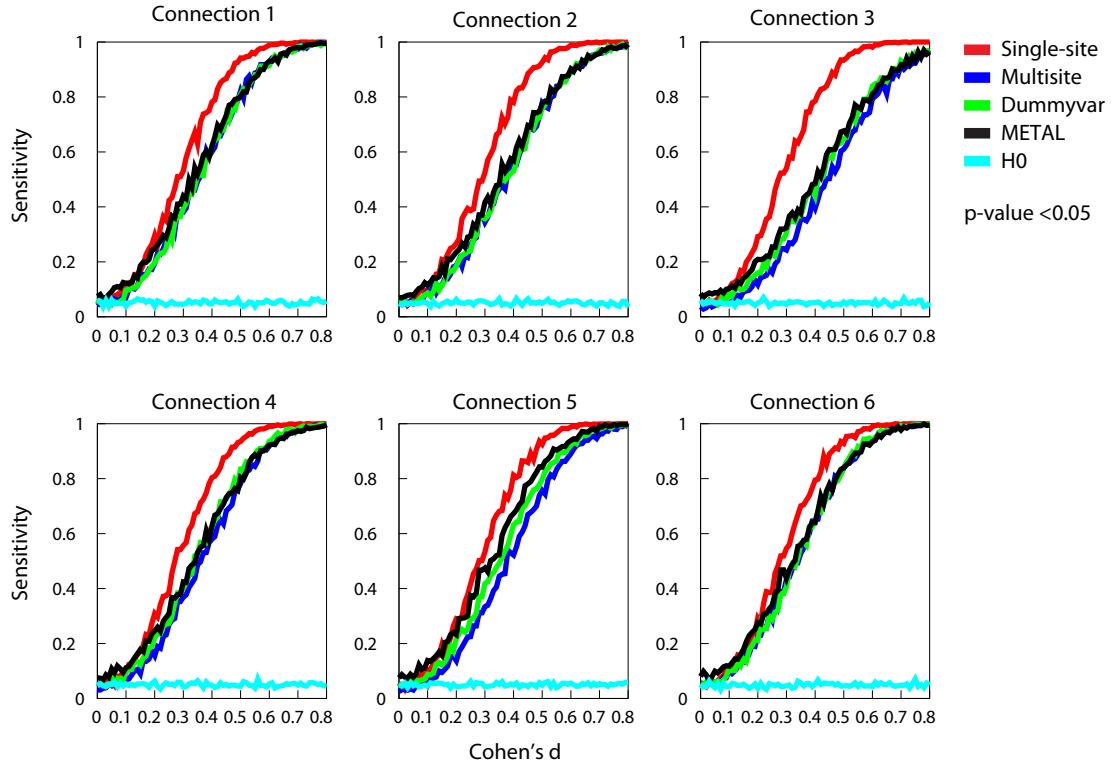


Figure 3.8: Power analysis for a resting-state fMRI study. A Monte-Carlo simulation was implemented to evaluate the power of a resting-state multi-site study. For each site and each sample, half of the subjects were randomly assigned to a 'treatment' group. For the subjects in this group, a value was 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.

869 **3.4 Conclusion**

870 We confirmed that multi-site acquisition introduce some variability in the dataset
871 although a single-site study with 200 subjects had only a marginally superior statistical
872 power than an analysis pooling 7 sites for an equivalent number of subjects, see Figure
873 3.7. In both cases, a high sensitivity (> 0.95) could be achieved for the effect size
874 observed by Goveas et al. (2011) which reported an effect equivalent to 1. We can
875 therefore conclude that it is feasible to acquire rs-fMRI data in across multiple sites and
876 correct for this topology.

FUTURE DEVELOPMENT879 **4.1 Large feature space**

880 Once the data is preprocessed properly, we need to obtain standard metric of con-
881 nectivity one of the most common metric is the Pearson correlation and is usually rep-
882 resented as a n x n connectivity matrix of every voxel in the gray matter (in the order of
883 10^4 voxels), there is an overwhelming number of 10^8 possible connections to examine
884 as a potential diagnostic tool and just a few example ($< 10^3$). We therefore have a very
885 large number of features, but only few of them are potentially relevant for predicting
886 the label. Even state-of-the-art classification algorithms (e.g. SVM (Cortes and Vapnik
887 1995)) cannot overcome the presence of large number of weakly relevant and redun-
888 dant features. This is usually attributed to "the curse of dimensionality" (Bellman and
889 Bellman 1961), or to the fact that irrelevant features decrease the ability of the learner
890 to discriminate between classes. Moreover, many machine-learning algorithms become
891 computationally intractable when the dimension is high. On the other hand once a re-
892 duce set of features has been chosen even the most basic classifiers can achieve high
893 performance classification.

894 In order to reduce that feature space, a number of solutions have been proposed that
895 take advantage of the underlying structure of the brain (van den Heuvel et al. 2009). Re-
896 gions are routinely defined using an anatomical parcellation (He et al. 2009), such as the
897 AAL template (Tzourio-Mazoyer et al. 2002). Anatomical parcels may however not well
898 match the organization of resting-state networks. We use a framework to generate data-
899 driven functional decomposition into resting-state networks based on the coherence of
900 BOLD activity at the individual or group level (Bellec 2006, Bellec et al. 2010b). When
901 a low number of networks (or scale) are used, this technique, called bootstrap analysis
902 of stable clusters (BASC), generates decompositions of the brain into distributed large-
903 scale networks, such as the default mode network (DMN). At high scales, it identifies

904 sub-networks and functional regions (Kelly et al. 2012). We can therefore use those
 905 parcellation units to reduce the feature space. Correlation is a good approach to asses
 906 connectivity but can be highly variable in noisy dataset this is why we would like to
 907 investigate another metric called "stability" based on evidence accumulation of cluster-
 908 ing on bootstrap samples, that could potentially be more consistent across subjects and
 909 scanning session resulting in improved prediction power and generalizability.

910 4.2 Feature selection and classification

911 A crucial point of our study is to identify the most discriminant features to use in
 912 the prediction model and account for various confounding factors (e.g. age, gender,
 913 education, multi-site effect and motion). This is future work that will be conducted in
 914 the next year and a half. In term of the classification tools, I'm considering the Linear
 915 Discriminant Analysis (LDA) because it is strait forward to include covariates in the
 916 model in order to account for confounding effects or SVM since It is a very popular
 917 method that should work well with the margin optimisation criteria that I plan to use for
 918 feature selection (Gilad-Bachrach et al. 2004, Kononenko et al. 1997). Although I do
 919 not plan to try a large set of classifiers and configurations since this will not be the focus
 920 of this project.

921 I will particularly focus on the stability of the features selected, this is an important
 922 point to obtain robust and consistent markers of the disease. In order to estimate stability
 923 of the selection I plan to use a Monte-Carlo estimation of the selection using bootstrap
 924 resampling (Bellec et al. 2010b, Efron and Tibshirani 1994) on the training dataset inside
 925 a 10-fold cross-validation. The large feature space that will be fed to this procedure will
 926 be a correlation matrix of network at multiple scales. As an example, let's take as an
 927 example 3 parcellations (e.g. scales 10, 50, 100) of the functional brains obtained from
 928 an independent dataset of normal subjects using the method describe in the previous
 929 section (Large feature space). Using the time-series of every parcellation I will obtain a
 930 correlation matrix A of size $R \times R$ and $R = 10 + 50 + 100 = 160$ resulting in a vectorized
 931 form of this matrix of size $L = R(R + 1)/2 = 12880$ unique features. The idea behind

932 this strategy is to capture interaction of small networks with larger networks instead of
933 just looking at the interaction of large network with large network or small network with
934 small network which could miss some important interaction. An example of previously
935 mention interaction of larger network with smaller ones is the known decrease in connec-
936 tivity between the hippocampal structure (small regions) with the default mode (a large
937 and extended network). This procedure could potentially capture those interactions that
938 would in term maximize the prediction accuracy as well as controlling for stability by
939 only retaining the most stable features identified by the feature selection therefore en-
940 forcing generalizability. Venkataraman et al. (2010) as used stability measures with Gini
941 importance metric and found in schizophrenia

942 An other approach would be to use ensemble techniques like AdaBoost (Freund and
943 Schapire 1997) to improve generalization performance. AdaBoost as been proved to be
944 remarkably resistant to overfitting (Schapire et al. 1998) The reality as depicted earlier
945 is that a lot of factors contribute to the noise and make such a classifier unfeasible. At
946 best, we can hope for classifiers to correctly classify the data most of the time and this is
947 what

948 **4.3 Industry application and translational effort**

949 I am involved in the industry (through my consulting work with the companies Neu-
950 roRX and Biospective), I'm advising on the fMRI analysis of multiple clinical trials and
951 looking at the feasibility of using fMRI in multicentric pharmaceutical trials (some of
952 these work have lead to publications). I'm also doing statistical analysis and propos-
953 ing biomarker tailored to the clinical questions of the pharmaceutical sponsor. I'm also
954 starting to get involved with the biomarker unit of the Canadian Consortium on Neu-
955 rodegeneration in Aging (CCNA) who wants to propose new biomarker for Alzheimer
956 disease. These efforts and collaborations are perfectly in line with the objectives of
957 my PhD. and are greatly contributing to translational findings and application of my re-
958 search in pharmaceutical trials as well as addressing concrete question in the field of
959 neuroimaging.

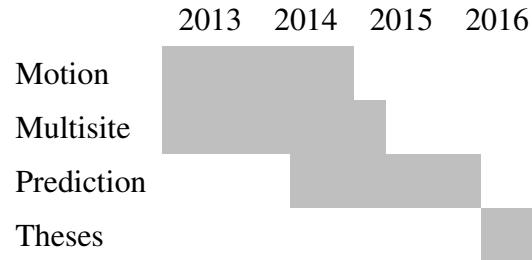
960 **4.4 Timeline**

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

- 962 ● Investigate motion impact on functional connectivity I'm currently writing a manuscript
963 that should be submitted at the end of this year (2014) on that specific topic. The
964 results have been presented (poster format) at the Organization for Human Brain
965 Mapping conference 2013 (OHBM) and at the Alzheimer's Association Interna-
966 tional Conference 2014 (AAIC).
- 967 ● Feasibility of multisite connectivity analysis (inter-site normalization) Most of the
968 analyses are completed and some of the results have been presented in two con-
969 ferences (poster format), namely OHBM2013 and AAIC2013. I'm planning to
970 submit the manuscript for this study mid 2015
- 971 ● Prediction
 - 972 – I'm currently experimenting with the prediction tools and I will start by test-
973 ing the pipeline on simple simulations.
 - 974 – Then on a real dataset I will try to predict the age of the subjects based on
975 their functional connectivity (this is known to be a large effect), I will use
976 for this the 1000 functional connectome dataset (350 adult subjects from
977 multiple sites).
 - 978 – Finally I will apply the pipeline on a dataset that I have compiled in the past
979 year, this dataset is composed of 313 elderly adults with and without cog-
980 nitive impairment of the Alzheimer type collected across 5 studies: ADNI2
981 study and 4 other studies based in Montreal, Canada, for a grand total of
982 126 CNE participants, 133 patients with MCI, and 54 patients with DAT. I
983 will try to classify the various population in a cross sectional analysis. I am
984 also planning to use a subset of the data namely the ADNI2 dataset who is
985 a longitudinal study to assesse the potential of the method to predict time of

986 conversion from pMCI to pDAT. I plan to be done with those analyses in the
987 winter of 2016 and publish the results in the following months.

- 988 • Finally I plan to submit my theses at the end of 2016



REFERENCES

- 990 A. Anand, Y. Li, Y. Wang, J. Wu, S. Gao, L. Bukhari, V. P. Mathews, A. Kalnin, and
991 M. J. Lowe. Activity and connectivity of brain mood regulating circuit in depression: a
992 functional magnetic resonance study. *Biological psychiatry*, 57(10):1079–1088, May
993 2005. ISSN 0006-3223. doi: 10.1016/j.biopsych.2005.02.021. URL <http://dx.doi.org/10.1016/j.biopsych.2005.02.021>.
- 995 Jessica R. Andrews-Hanna, Jay S. Reidler, Jorge Sepulcre, Renee Poulin, and Randy L.
996 Buckner. Functional-anatomic fractionation of the brain’s default network. *Neuron*,
997 65(4):550–562, February 2010. ISSN 1097-4199. doi: 10.1016/j.neuron.2010.02.005.
998 URL <http://dx.doi.org/10.1016/j.neuron.2010.02.005>.
- 999 R. Baumgartner, C. Windischberger, and E. Moser. Quantification in Func-
1000 tional Magnetic Resonance Imaging: Fuzzy Clustering vs. Correlation Anal-
1001 ysis. *Magnetic Resonance Imaging*, 16(2):115–125, February 1998. doi:
1002 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).
- 1004 Yashar Behzadi, Khaled Restom, Joy Liau, and Thomas T. Liu. A component based
1005 noise correction method (CompCor) for BOLD and perfusion based fMRI. *NeuroIm-
1006 age*, 37(1):90–101, August 2007. ISSN 1053-8119. doi: 10.1016/j.neuroimage.2007.
1007 04.042. URL <http://dx.doi.org/10.1016/j.neuroimage.2007.04.042>.
- 1009 P. Bellec, F. Carbonell, V. Perlberg, and A. C. Evans. *A neuroimaging analysis kit for
1010 Octave and Matlab*, 2010a. URL <http://code.google.com/p/niak/>.
- 1011 Pierre Bellec. *Longitudinal study of large-scale networks in the human brain using
1012 fMRI : Methods and application to motor skill learning*. PhD thesis, Université Paris
1013 XI, Orsay, Paris, February 2006. URL http://exocet.imec.jussieu.fr/~pbellec/data/these_bellec.pdf.

- 1015 Pierre Bellec, Pedro Rosa-Neto, Oliver C. Lyttelton, Habib Benali, and Alan C. Evans.
1016 Multi-level bootstrap analysis of stable clusters in resting-state fMRI. *NeuroImage*,
1017 51(3):1126–1139, July 2010b. ISSN 1095-9572. doi: 10.1016/j.neuroimage.2010.
1018 02.082. URL <http://dx.doi.org/10.1016/j.neuroimage.2010.02.082>.
1019
- 1020 R. Bellman and R.E. Bellman. *Adaptive Control Processes: A Guided Tour*. 'Rand
1021 Corporation. Research studies. Princeton University Press, 1961. URL <http://books.google.ca/books?id=POAmAAAAMAAJ>.
- 1023 B. Biswal, F. Z. Yetkin, V. M. Haughton, and J. S. Hyde. Functional connectivity in the
1024 motor cortex of resting human brain using echo-planar MRI. *Magnetic resonance in*
1025 *medicine*, 34(4):537–541, October 1995. ISSN 0740-3194. URL <http://view.ncbi.nlm.nih.gov/pubmed/8524021>.
- 1027 Bharat B. Biswal, Maarten Mennes, Xi-Nian N. Zuo, Suril Gohel, Clare Kelly, Steve M.
1028 Smith, Christian F. Beckmann, Jonathan S. Adelstein, Randy L. Buckner, Stan Col-
1029 combe, Anne-Marie M. Dogonowski, Monique Ernst, Damien Fair, Michelle Hamp-
1030 son, Matthew J. Hoptman, James S. Hyde, Vesa J. Kiviniemi, Rolf Kötter, Shi-Jiang J.
1031 Li, Ching-Po P. Lin, Mark J. Lowe, Clare Mackay, David J. Madden, Kristoffer H.
1032 Madsen, Daniel S. Margulies, Helen S. Mayberg, Katie McMahon, Christopher S.
1033 Monk, Stewart H. Mostofsky, Bonnie J. Nagel, James J. Pekar, Scott J. Peltier,
1034 Steven E. Petersen, Valentin Riedl, Serge A. Rombouts, Bart Rypma, Bradley L.
1035 Schlaggar, Sein Schmidt, Rachael D. Seidler, Greg J. Siegle, Christian Sorg, Gao-
1036 Jun J. Teng, Juha Veijola, Arno Villringer, Martin Walter, Lihong Wang, Xu-Chu C.
1037 Weng, Susan Whitfield-Gabrieli, Peter Williamson, Christian Windischberger, Yu-
1038 Feng F. Zang, Hong-Ying Y. Zhang, F. Xavier Castellanos, and Michael P. Mil-
1039 ham. Toward discovery science of human brain function. *Proceedings of the Na-*
1040 *tional Academy of Sciences of the United States of America*, 107(10):4734–4739,
1041 March 2010. ISSN 1091-6490. doi: 10.1073/pnas.0911855107. URL <http://dx.doi.org/10.1073/pnas.0911855107>.
1042

- 1043 Randy L. Buckner, Jessica R. Andrews-Hanna, and Daniel L. Schacter. The Brain's
1044 Default Network. *Annals of the New York Academy of Sciences*, 1124(1):1–38, March
1045 2008. ISSN 0077-8923. doi: 10.1196/annals.1440.011. URL <http://dx.doi.org/10.1196/annals.1440.011>.
- 1047 Randy L. Buckner, Jorge Sepulcre, Tanveer Talukdar, Fenna M. Krienen, Hesheng
1048 Liu, Trey Hedden, Jessica R. Andrews-Hanna, Reisa A. Sperling, and Keith A.
1049 Johnson. Cortical hubs revealed by intrinsic functional connectivity: mapping, as-
1050 sessment of stability, and relation to Alzheimer's disease. *The Journal of neu-*
1051 *roscience : the official journal of the Society for Neuroscience*, 29(6):1860–1873,
1052 February 2009. ISSN 1529-2401. doi: 10.1523/JNEUROSCI.5062-08.2009. URL
1053 <http://dx.doi.org/10.1523/JNEUROSCI.5062-08.2009>.
- 1054 F. Carbonell, P. Bellec, and A. Shmuel. Quantification of the impact of a confounding
1055 variable on functional connectivity confirms anti-correlated networks in the resting-
1056 state. *Neuroimage*, 86:343–353, Feb 2014. doi: 10.1016/j.neuroimage.2013.10.013.
1057 URL <http://dx.doi.org/10.1016/j.neuroimage.2013.10.013>.
- 1058 Gang Chen, B. Douglas Ward, Chunming Xie, Wenjun Li, Zhilin Wu, Jennifer L. Jones,
1059 Malgorzata Franczak, Piero Antuono, and Shi-Jiang Li. Classification of Alzheimer
1060 Disease, Mild Cognitive Impairment, and Normal Cognitive Status with Large-Scale
1061 Network Analysis Based on Resting-State Functional MR Imaging. *Radiology*, 259
1062 (1):213–221, April 2011. doi: 10.1148/radiol.10100734. URL <http://dx.doi.org/10.1148/radiol.10100734>.
- 1064 G. Chetelat, B. Desgranges, V. de la Sayette, F. Viader, F. Eustache, and J-C. Baron. Mild
1065 cognitive impairment: Can fdg-pet predict who is to rapidly convert to alzheimer's
1066 disease? *Neurology*, 60(8):1374–1377, Apr 2003.
- 1067 Z. H. Cho, S. C. Chung, D. W. Lim, and E. K. Wong. Effects of the acoustic noise of
1068 the gradient systems on fmri: a study on auditory, motor, and visual cortices. *Magn
1069 Reson Med*, 39(2):331–335, Feb 1998.

- 1070 D. L. Collins, A. P. Zijdenbos, V. Kollokian, J. G. Sled, N. J. Kabani, C. J. Holmes,
1071 and A. C. Evans. Design and construction of a realistic digital brain phantom. *IEEE*
1072 *Trans Med Imaging*, 17(3):463–468, June 1998. ISSN 0278-0062. URL <http://view.ncbi.nlm.nih.gov/pubmed/9735909>.
- 1074 Dietmar Cordes, Victor M. Haughton, Konstantinos Arfanakis, Gary J. Wendt, Patrick A.
1075 Turski, Chad H. Moritz, Michelle A. Quigley, and M. Elizabeth Meyerand. Mapping
1076 Functionally Related Regions of Brain with Functional Connectivity MR Imaging.
1077 *American Journal of Neuroradiology*, 21(9):1636–1644, October 2000. URL <http://www.ajnr.org/content/21/9/1636.abstract>.
- 1079 Dietmar Cordes, Victor M. Haughton, Konstantinos Arfanakis, John D. Carew,
1080 Patrick A. Turski, Chad H. Moritz, Michelle A. Quigley, and Elizabeth M. Meyerand.
1081 Frequencies Contributing to Functional Connectivity in the Cerebral Cortex in
1082 "Resting-state" Data. *AJNR Am J Neuroradiol*, 22(7):1326–1333, August 2001. URL
1083 <http://www.ajnr.org/cgi/content/abstract/22/7/1326>.
- 1084 Dietmar Cordes, Vic Haughton, John D. Carew, Konstantinos Arfanakis, and Ken
1085 Maravilla. Hierarchical clustering to measure connectivity in fMRI resting-state
1086 data. *Magnetic Resonance Imaging*, 20(4):305–317, May 2002. ISSN 0730725X.
1087 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).
- 1089 Corinna Cortes and Vladimir Vapnik. Support-vector networks. *Machine Learning*, 20
1090 (3):273–297, September 1995. ISSN 0885-6125. doi: 10.1007/BF00994018. URL
1091 <http://dx.doi.org/10.1007/BF00994018>.
- 1092 Sergi G Costafreda, Akash Khanna, Janaina Mourao-Miranda, and Cynthia H Y
1093 Fu. Neural correlates of sad faces predict clinical remission to cognitive be-
1094 havioural therapy in depression. *Neuroreport*, 20(7):637–641, May 2009. doi: 10.
1095 1097/WNR.0b013e3283294159. URL <http://dx.doi.org/10.1097/WNR.0b013e3283294159>.

- 1097 Kathryn R Cullen, Dylan G Gee, Bonnie Klimes-Dougan, Vilma Gabbay, Leslie Hul-
1098 vershorn, Bryon A Mueller, Jazmin Camchong, Christopher J Bell, Alaa Houry, San-
1099 jiv Kumra, Kelvin O Lim, F. Xavier Castellanos, and Michael P Milham. A pre-
1100 liminary study of functional connectivity in comorbid adolescent depression. *Neu-*
1101 *rosci Lett*, 460(3):227–231, Sep 2009. doi: 10.1016/j.neulet.2009.05.022. URL
1102 <http://dx.doi.org/10.1016/j.neulet.2009.05.022>.
- 1103 Zhengjia Dai, Chaogan Yan, Kuncheng Li, Zhiqun Wang, Jinhui Wang, Miao Cao, Qix-
1104 iang Lin, Ni Shu, Mingrui Xia, Yanchao Bi, and Yong He. Identifying and mapping
1105 connectivity patterns of brain network hubs in alzheimer’s disease. *Cereb Cortex*,
1106 Oct 2014. doi: 10.1093/cercor/bhu246. URL <http://dx.doi.org/10.1093/cercor/bhu246>.
- 1108 Marcello D’Amelio and Paolo Maria Rossini. Brain excitability and connectivity of
1109 neuronal assemblies in alzheimer’s disease: from animal models to human findings.
1110 *Progress in neurobiology*, 99(1):42–60, October 2012. ISSN 1873-5118. URL <http://view.ncbi.nlm.nih.gov/pubmed/22789698>.
- 1112 J. S. Damoiseaux, S. A. R. B. Rombouts, F. Barkhof, P. Scheltens, C. J. Stam, S. M.
1113 Smith, and C. F. Beckmann. Consistent resting-state networks across healthy subjects.
1114 *Proceedings of the National Academy of Sciences*, 103(37):13848–13853, September
1115 2006. ISSN 1091-6490. doi: 10.1073/pnas.0601417103. URL <http://dx.doi.org/10.1073/pnas.0601417103>.
- 1117 Jessica S. Damoiseaux, Katherine E. Prater, Bruce L. Miller, and Michael D. Gre-
1118 icius. Functional connectivity tracks clinical deterioration in Alzheimer’s disease.
1119 *Neurobiology of Aging*, 33(4):828.e19–828.e30, April 2012. ISSN 01974580. doi:
1120 10.1016/j.neurobiolaging.2011.06.024. URL <http://dx.doi.org/10.1016/j.neurobiolaging.2011.06.024>.
- 1122 Christian Dansereau, Celine Risterucci, Emilio Merlo Pich, Douglas Arnold, and
1123 Pierre Bellec. A power analysis for multisite studies in resting-state functional

- 1124 connectivity, with an application to clinical trials in alzheimer's disease. volume
1125 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
1126 Alzheimer's Association International Conference 2013.
1127
1128
- 1129 Christian Dansereau, Pierre Orban, and Pierre Bellec. Impact of motion on resting-
1130 state fmri connectivity in healthy aging and alzheimer's disease, and possible reme-
1131 dies. volume 10, pages P49 –, 2014. 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
1132 Alzheimer's Association International Conference 2014.
1133
1134
- 1135 John Desmond and Gary Glover. Estimating sample size in functional mri (fmri) neu-
1136 roimaging studies: Statistical power analyses. *Journal of Neuroscience Methods*, 118
1137 (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).
1138
- 1139 V. Edward, C. Windischberger, R. Cunnington, M. Erdler, R. Lanzenberger, D. Mayer,
1140 W. Endl, and R. Beisteiner. Quantification of fmri artifact reduction by a novel plaster
1141 cast head holder. *Hum Brain Mapp*, 11(3):207–213, Nov 2000.
- 1142 Bradley Efron and R. J. Tibshirani. *An Introduction to the Bootstrap (Chapman &*
- 1143 Hall/CRC Monographs on Statistics & Applied Probability). Chapman and Hall/CRC,
1144 1 edition, May 1994. ISBN 0412042312. URL <http://www.worldcat.org/isbn/0412042312>.
1145
- 1146 M. R. Elliott, R. W. Bowtell, and P. G. Morris. The effect of scanner sound in visual,
1147 motor, and auditory functional mri. *Magn Reson Med*, 41(6):1230–1235, Jun 1999.
- 1148 Damien A Fair, Joel T Nigg, Swathi Iyer, Deepti Bathula, Kathryn L Mills, Nico U F
1149 Dosenbach, Bradley L Schlaggar, Maarten Mennes, David Gutman, Saroja Bangaru,
1150 Jan K Buitelaar, Daniel P Dickstein, Adriana Di Martino, David N Kennedy, Clare

- 1151 Kelly, Beatriz Luna, Julie B Schweitzer, Katerina Velanova, Yu-Feng Wang, Stew-
1152 art Mostofsky, F. Xavier Castellanos, and Michael P Milham. Distinct neural signa-
1153 tures detected for adhd subtypes after controlling for micro-movements in resting state
1154 functional connectivity mri data. *Front Syst Neurosci*, 6:80, 2012. doi: 10.3389/fnsys.
1155 2012.00080. URL <http://dx.doi.org/10.3389/fnsys.2012.00080>.
- 1156 Yong Fan, Dinggang Shen, and Christos Davatzikos. Classification of structural images
1157 via high-dimensional image warping, robust feature extraction, and svm. *Med Image
1158 Comput Comput Assist Interv*, 8(Pt 1):1–8, 2005.
- 1159 Fang Fang, Scott O Murray, and Sheng He. Duration-dependent fmri adaptation and
1160 distributed viewer-centered face representation in human visual cortex. *Cereb Cortex*,
1161 17(6):1402–1411, Jun 2007. doi: 10.1093/cercor/bhl053. URL <http://dx.doi.org/10.1093/cercor/bhl053>.
- 1163 David A Feinberg, Steen Moeller, Stephen M Smith, Edward Auerbach, Sudhir Ra-
1164 manna, Matthias Gunther, Matt F Glasser, Karla L Miller, Kamil Ugurbil, and Essa
1165 Yacoub. Multiplexed echo planar imaging for sub-second whole brain fmri and
1166 fast diffusion imaging. *PLoS One*, 5(12):e15710, 2010. doi: 10.1371/journal.pone.
1167 0015710. URL <http://dx.doi.org/10.1371/journal.pone.0015710>.
- 1168 Vladimir Fonov, Alan C. Evans, Kelly Botteron, C. Robert Almli, Robert C. McKinstry,
1169 D. Louis Collins, and Brain Development Cooperative Group. Unbiased average age-
1170 appropriate atlases for pediatric studies. *NeuroImage*, 54(1):313–327, January 2011.
1171 ISSN 1095-9572. doi: 10.1016/j.neuroimage.2010.07.033. URL <http://dx.doi.org/10.1016/j.neuroimage.2010.07.033>.
- 1173 Michael D. Fox and Marcus E. Raichle. Spontaneous fluctuations in brain activ-
1174 ity observed with functional magnetic resonance imaging. *Nat Rev Neurosci*, 8
1175 (9):700–711, September 2007. ISSN 1471-003X. doi: 10.1038/nrn2201. URL
1176 <http://dx.doi.org/10.1038/nrn2201>.
- 1177 Michael D Fox, Dongyang Zhang, Abraham Z Snyder, and Marcus E Raichle. The global

- 1178 signal and observed anticorrelated resting state brain networks. *J Neurophysiol*, 101
1179 (6):3270–3283, Jun 2009. doi: 10.1152/jn.90777.2008. URL <http://dx.doi.org/10.1152/jn.90777.2008>.
- 1181 Yoav Freund and Robert E Schapire. A decision-theoretic generalization of on-line
1182 learning and an application to boosting. *Journal of Computer and System Sciences*,
1183 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>.
- 1186 Lee Friedman and Gary Glover. Report on a multicenter fmri quality assurance protocol.
1187 *Journal of magnetic resonance imaging : JMRI*, 23(6):827–839, June 2006. ISSN
1188 1053-1807. URL <http://dx.doi.org/10.1002/jmri.20583>.
- 1189 Lee Friedman, Gary Glover, and The FBIRN Consortium. Reducing interscanner vari-
1190 ability of activation in a multicenter fmri study: Controlling for signal-to-fluctuation-
1191 noise-ratio (sfnr) differences. *NeuroImage*, 33(2):471–481, November 2006. ISSN
1192 10538119. URL <http://dx.doi.org/10.1016/j.neuroimage.2006.07.012>.
- 1194 Cynthia H Y Fu, Janaina Mourao-Miranda, Sergi G Costafreda, Akash Khanna, Andre F
1195 Marquand, Steve C R Williams, and Michael J Brammer. Pattern classification of sad
1196 facial processing: toward the development of neurobiological markers in depression.
1197 *Biol Psychiatry*, 63(7):656–662, Apr 2008. doi: 10.1016/j.biopsych.2007.08.020.
1198 URL <http://dx.doi.org/10.1016/j.biopsych.2007.08.020>.
- 1199 Ran Gilad-Bachrach, Amir Navot, and Naftali Tishby. Margin based feature selection -
1200 theory and algorithms. In *Proceedings of the Twenty-first International Conference on*
1201 *Machine Learning*, ICML '04, pages 43–, New York, NY, USA, 2004. ACM. ISBN
1202 1-58113-838-5. doi: 10.1145/1015330.1015352. URL <http://doi.acm.org/10.1145/1015330.1015352>.
- 1204 Federico Giove, Tommaso Gili, Vittorio Iacobella, Emiliano Macaluso, and Bruno Mar-

- 1205 aviglia. Images-based suppression of unwanted global signals in resting-state func-
1206 tional connectivity studies. *Magnetic resonance imaging*, 27(8):1058–1064, October
1207 2009. ISSN 1873-5894. doi: 10.1016/j.mri.2009.06.004. URL <http://dx.doi.org/10.1016/j.mri.2009.06.004>.
- 1209 Joseph S. Goveas, Chunming Xie, B. Douglas Ward, Zhilin Wu, Wenjun Li, Malgo-
1210 rzata Franczak, Jennifer L. Jones, Piero G. Antuono, and Shi-Jiang J. Li. Recovery
1211 of hippocampal network connectivity correlates with cognitive improvement in mild
1212 Alzheimer’s disease patients treated with donepezil assessed by resting-state fMRI.
1213 *Journal of magnetic resonance imaging : JMRI*, 34(4):764–773, October 2011. ISSN
1214 1522-2586. doi: 10.1002/jmri.22662. URL <http://dx.doi.org/10.1002/jmri.22662>.
- 1216 M. D. Greicius, B. Krasnow, A. L. Reiss, and V. Menon. Functional connectivity in
1217 the resting brain: A network analysis of the default mode hypothesis. *Proc. Natl.
1218 Acad. Sci.*, 100(1):253–258, January 2003. ISSN 1091-6490. doi: 10.1073/pnas.
1219 0135058100. URL <http://dx.doi.org/10.1073/pnas.0135058100>.
- 1220 Michael D. Greicius, Gaurav Srivastava, Allan L. Reiss, and Vinod Menon. Default-
1221 mode network activity distinguishes Alzheimer’s disease from healthy aging: Evi-
1222 dence from functional MRI. *Proceedings of the National Academy of Sciences of
1223 the United States of America*, 101(13):4637–4642, March 2004. ISSN 1091-6490.
1224 doi: 10.1073/pnas.0308627101. URL <http://dx.doi.org/10.1073/pnas.0308627101>.
- 1226 Michael D. Greicius, Benjamin H. Flores, Vinod Menon, Gary H. Glover, Hugh B.
1227 Solvason, Heather Kenna, Allan L. Reiss, and Alan F. Schatzberg. Resting-state
1228 functional connectivity in major depression: abnormally increased contributions from
1229 subgenual cingulate cortex and thalamus. *Biological psychiatry*, 62(5):429–437,
1230 September 2007. ISSN 0006-3223. doi: 10.1016/j.biopsych.2006.09.020. URL
1231 <http://dx.doi.org/10.1016/j.biopsych.2006.09.020>.
- 1232 Michael D. Greicius, Kaustubh Supekar, Vinod Menon, and Robert F. Dougherty.

- 1233 Resting-state functional connectivity reflects structural connectivity in the default
1234 mode network. *Cerebral cortex (New York, N.Y. : 1991)*, 19(1):72–78, January 2009.
1235 ISSN 1460-2199. doi: 10.1093/cercor/bhn059. URL <http://dx.doi.org/10.1093/cercor/bhn059>.
- 1237 D. A. Gusnard and M. E. Raichle. Searching for a baseline: Functional imaging and the
1238 resting human brain. *Nature Reviews Neuroscience*, 2(10):685–694, October 2001.
1239 ISSN 1471-003X. doi: 10.1038/35094500. URL <http://dx.doi.org/10.1038/35094500>.
- 1241 Tim Hahn, Andre F Marquand, Ann-Christine Ehli, Thomas Dresler, Sarah Kittel-
1242 Schneider, Tomasz A Jarczok, Klaus-Peter Lesch, Peter M Jakob, Janaina Mourao-
1243 Miranda, Michael J Brammer, and Andreas J Fallgatter. Integrating neurobiological
1244 markers of depression. *Arch Gen Psychiatry*, 68(4):361–368, Apr 2011. doi:
1245 10.1001/archgenpsychiatry.2010.178. URL <http://dx.doi.org/10.1001/archgenpsychiatry.2010.178>.
- 1247 M. Hampson, B. S. Peterson, P. Skudlarski, J. C. Gatenby, and J. C. Gore. Detection of
1248 functional connectivity using temporal correlations in MR images. *Hum Brain Mapp*,
1249 15(4):247–262, April 2002. ISSN 1065-9471. URL <http://view.ncbi.nlm.nih.gov/pubmed/11835612>.
- 1251 John Hardy and Dennis J Selkoe. The amyloid hypothesis of alzheimer’s disease:
1252 progress and problems on the road to therapeutics. *Science*, 297(5580):353–356, Jul
1253 2002. doi: 10.1126/science.1072994. URL <http://dx.doi.org/10.1126/science.1072994>.
- 1255 Egbert Hartstra, Simone KÃaijn, Tom Verguts, and Marcel Brass. The implementation
1256 of verbal instructions: an fmri study. *Hum Brain Mapp*, 32(11):1811–1824, Nov 2011.
1257 doi: 10.1002/hbm.21152. URL <http://dx.doi.org/10.1002/hbm.21152>.
- 1258 Yong He, Zhang Chen, Gaolang Gong, and Alan Evans. Neuronal networks in
1259 Alzheimer’s disease. *The Neuroscientist : a review journal bringing neuro-*

- 1260 *biology, neurology and psychiatry*, 15(4):333–350, August 2009. ISSN 1073-
1261 8584. doi: 10.1177/1073858409334423. URL <http://dx.doi.org/10.1177/1073858409334423>.
- 1263 David J Heeger and David Ress. What does fmri tell us about neuronal activity? *Nat
1264 Rev Neurosci*, 3(2):142–151, Feb 2002. doi: 10.1038/nrn730. URL <http://dx.doi.org/10.1038/nrn730>.
- 1266 R. D. Hoge, J. Atkinson, B. Gill, G. R. Crelier, S. Marrett, and G. B. Pike. Investigation
1267 of bold signal dependence on cerebral blood flow and oxygen consumption: the
1268 deoxyhemoglobin dilution model. *Magn Reson Med*, 42(5):849–863, Nov 1999.
- 1269 James X. Hu, Hongyu Zhao, and Harrison H. Zhou. False Discovery Rate Con-
1270 trol With Groups. *Journal of the American Statistical Association*, 105(491):1215–
1271 1227, September 2010. ISSN 0162-1459. doi: 10.1198/jasa.2010.tm09329. URL
1272 <http://dx.doi.org/10.1198/jasa.2010.tm09329>.
- 1273 C. R. Jack, R. C. Petersen, Y. C. Xu, P. C. O’Brien, G. E. Smith, R. J. Ivnik, B. F.
1274 Boeve, S. C. Waring, E. G. Tangalos, and E. Kokmen. Prediction of ad with mri-based
1275 hippocampal volume in mild cognitive impairment. *Neurology*, 52(7):1397–1403, Apr
1276 1999.
- 1277 Clifford R Jack, David S Knopman, William J Jagust, Leslie M Shaw, Paul S Aisen,
1278 Michael W Weiner, Ronald C Petersen, and John Q Trojanowski. Hypothetical model
1279 of dynamic biomarkers of the alzheimer’s pathological cascade. *Lancet Neurol*, 9
1280 (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).
- 1282 Jaeseung Jeong. Eeg dynamics in patients with alzheimer’s disease. *Clin Neurophysiol*,
1283 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>.
- 1285 Tianzi Jiang, Yong He, Yufeng Zang, and Xuchu Weng. Modulation of functional
1286 connectivity during the resting state and the motor task. *Hum. Brain Mapp.*, 22

- 1287 (1):63–71, May 2004. ISSN 1065-9471. doi: 10.1002/hbm.20012. URL <http://dx.doi.org/10.1002/hbm.20012>.
- 1288
- 1289 Xi Jiang, Xin Zhang, and Dajiang Zhu. Intrinsic functional component analysis via
1290 sparse representation on alzheimer’s disease neuroimaging initiative database. *Brain
1291 Connect*, 4(8):575–586, Oct 2014. doi: 10.1089/brain.2013.0221. URL [http://
1292 dx.doi.org/10.1089/brain.2013.0221](http://dx.doi.org/10.1089/brain.2013.0221).
- 1293 Yasuhiro Kawasaki, Michio Suzuki, Ferath Kherif, Tsutomu Takahashi, Shi-Yu Zhou,
1294 Kazue Nakamura, Mie Matsui, Tomiki Sumiyoshi, Hikaru Seto, and Masayoshi Ku-
1295 rachi. Multivariate voxel-based morphometry successfully differentiates schizophre-
1296 nia patients from healthy controls. *Neuroimage*, 34(1):235–242, Jan 2007. doi:
1297 10.1016/j.neuroimage.2006.08.018. URL [http://dx.doi.org/10.1016/j.
neuroimage.2006.08.018](http://dx.doi.org/10.1016/j.
1298 neuroimage.2006.08.018).
- 1299 Clare Kelly, Roberto Toro, Adriana Di Martino, Christine L. Cox, Pierre Bellec,
1300 F. Xavier Castellanos, and Michael P. Milham. A convergent functional architec-
1301 ture of the insula emerges across imaging modalities. *NeuroImage*, 61(4):1129–
1302 1142, July 2012. ISSN 10538119. doi: 10.1016/j.neuroimage.2012.03.021. URL
1303 [http://dx.doi.org/10.1016/j.
neuroimage.2012.03.021](http://dx.doi.org/10.1016/j.neuroimage.2012.03.021).
- 1304 Markus Klärhofer, Markus Barth, and Ewald Moser. Comparison of multi-echo spiral
1305 and echo planar imaging in functional mri. *Magn Reson Imaging*, 20(4):359–364,
1306 May 2002.
- 1307 Igor Kononenko, Edvard Simec, and Marko Robnik-Sikonja. Overcoming the myopia
1308 of inductive learning algorithms with relieff. *Applied Intelligence*, 7:39–55, 1997.
- 1309 S. B. Kotsiantis. Supervised machine learning: A review of classification tech-
1310 niques. pages 3–24, 2007. URL <http://dl.acm.org/citation.cfm?id=1566770.1566773>.
- 1311
- 1312 Zhiqiang Lao, Dinggang Shen, Zhong Xue, Bilge Karacali, Susan M Resnick, and Chris-

- tos 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, Zhenning 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>.

- 1340 Daniel S. Margulies, Justin L. Vincent, Clare Kelly, Gabriele Lohmann, Lucina Q. Ud-
1341 din, Bharat B. Biswal, Arno Villringer, F. Xavier Castellanos, Michael P. Milham, and
1342 Michael Petrides. Precuneus shares intrinsic functional architecture in humans and
1343 monkeys. *Proceedings of the National Academy of Sciences of the United States of*
1344 *America*, November 2009. ISSN 1091-6490. doi: 10.1073/pnas.0905314106. URL
1345 <http://dx.doi.org/10.1073/pnas.0905314106>.
- 1346 Andre F Marquand, Janaina Mourão-Miranda, Michael J Brammer, Anthony J Cleare,
1347 and Cynthia H Y Fu. Neuroanatomy of verbal working memory as a diagnostic
1348 biomarker for depression. *Neuroreport*, 19(15):1507–1511, Oct 2008. doi: 10.
1349 1097/WNR.0b013e328310425e. URL <http://dx.doi.org/10.1097/WNR.0b013e328310425e>.
- 1351 Niklas Mattsson, Henrik Zetterberg, Oskar Hansson, Niels Andreasen, Lucilla Parnetti,
1352 Michael Jonsson, Sanna-Kaisa Herukka, Wiesje M van der Flier, Marinus A Blanken-
1353 stein, Michael Ewers, Kenneth Rich, Elmar Kaiser, Marcel Verbeek, Magda Tsolaki,
1354 Ezra Mulugeta, Erik Rosén, Dag Aarsland, Pieter Jelle Visser, Johannes Schröder,
1355 Jan Marcusson, Mony de Leon, Harald Hampel, Philip Scheltens, Tuula Pirttilä,
1356 Anders Wallin, Maria Eriksdotter Jähnigen, Lennart Minthon, Bengt Winblad, and
1357 Kaj Blennow. Csf biomarkers and incipient alzheimer disease in patients with mild
1358 cognitive impairment. *JAMA*, 302(4):385–393, Jul 2009. doi: 10.1001/jama.2009.
1359 1064. URL <http://dx.doi.org/10.1001/jama.2009.1064>.
- 1360 M. J. McKeown, S. Makeig, G. G. Brown, T. P. Jung, S. S. Kindermann, A. J. Bell,
1361 and T. J. Sejnowski. Analysis of fMRI data by blind separation into independent
1362 spatial components. *Hum Brain Mapp*, 6(3):160–188, 1998. ISSN 1065-9471. URL
1363 <http://view.ncbi.nlm.nih.gov/pubmed/9673671>.
- 1364 G. McKhann, D. Drachman, M. Folstein, R. Katzman, D. Price, and E. M. Stadlan.
1365 Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work
1366 Group under the auspices of Department of Health and Human Services Task Force on

- 1367 Alzheimer's Disease. *Neurology*, 34(7):939–944, July 1984. ISSN 0028-3878. URL
1368 <http://www.neurology.org/content/34/7/939.abstract>.
- 1369 V. Menon, K.O. Lim, J.H. Anderson, J. Johnson, and A. Pfefferbaum. Design and ef-
1370 ficacy of a head-coil bite bar for reducing movement-related artifacts during func-
1371 tional mri scanning. *Behavior Research Methods, Instruments, & Computers*, 29
1372 (4):589–594, 1997. ISSN 0743-3808. doi: 10.3758/BF03210613. URL <http://dx.doi.org/10.3758/BF03210613>.
- 1374 A. Meyer-Baese, A. Wismueller, and O. Lange. Comparison of two exploratory data
1375 analysis methods for fMRI: unsupervised clustering versus independent component
1376 analysis. *IEEE Trans Inf Technol Biomed*, 8(3):387–398, September 2004. ISSN
1377 1089-7771. URL <http://view.ncbi.nlm.nih.gov/pubmed/15484444>.
- 1378 Michal Mikl, Radek Marecek, Petr Hlustík, Martina Pavlicová, Ales Drastich, Pavel
1379 Chlebus, Milan Brázdil, and Petr Krupa. Effects of spatial smoothing on fmri group
1380 inferences. *Magnetic resonance imaging*, 26(4):490–503, May 2008. ISSN 0730-
1381 725X. URL <http://view.ncbi.nlm.nih.gov/pubmed/18060720>.
- 1382 John C Morris. Early-stage and preclinical alzheimer disease. *Alzheimer Dis Assoc
1383 Disord*, 19(3):163–165, 2005.
- 1384 Janaina Mourao-Miranda, Arun L W Bokde, Christine Born, Harald Hampel, and Martin
1385 Stetter. Classifying brain states and determining the discriminating activation patterns:
1386 Support vector machine on functional mri data. *Neuroimage*, 28(4):980–995, Dec
1387 2005. doi: 10.1016/j.neuroimage.2005.06.070. URL <http://dx.doi.org/10.1016/j.neuroimage.2005.06.070>.
- 1389 Susanne G Mueller, Michael W Weiner, Leon J Thal, Ronald C Petersen, Clifford
1390 Jack, William Jagust, John Q Trojanowski, Arthur W Toga, and Laurel Beckett. The
1391 alzheimer's disease neuroimaging initiative. *Neuroimaging Clin N Am*, 15(4):869–77,
1392 xi–xii, Nov 2005. doi: 10.1016/j.nic.2005.09.008. URL <http://dx.doi.org/10.1016/j.nic.2005.09.008>.

- 1394 Kevin Murphy, Rasmus M. Birn, Daniel A. Handwerker, Tyler B. Jones, and Peter A.
1395 Bandettini. The impact of global signal regression on resting state correlations: are
1396 anti-correlated networks introduced? *NeuroImage*, 44(3):893–905, February 2009.
1397 ISSN 1095-9572. doi: 10.1016/j.neuroimage.2008.09.036. URL <http://dx.doi.org/10.1016/j.neuroimage.2008.09.036>.
- 1398
- 1399 Jared Nielsen, Brandon Zielinski, Thomas Fletcher, Andrew Alexander, Nicholas Lange,
1400 Erin Bigler, Janet Lainhart, and Jeffrey Anderson. Multisite functional connectivity
1401 mri classification of autism: Abide results. *Frontiers in human neuroscience*, 7:–,
1402 2013. ISSN 1662-5161. URL <http://view.ncbi.nlm.nih.gov/pubmed/24093016>.
- 1403
- 1404 Ilia Nouretdinov, Sergi Costafreda, Alexander Gammerman, Alexey Chervonenkis,
1405 Vladimir Vovk, Vladimir Vapnik, and Cynthia Fu. Machine learning classification
1406 with confidence: Application of transductive conformal predictors to mri-based diag-
1407 nostic and prognostic markers in depression. *NeuroImage*, 56(2):809–813, May 2011.
1408 ISSN 10538119. URL <http://dx.doi.org/10.1016/j.neuroimage.2010.05.023>.
- 1409
- 1410 S. Ogawa, T. M. Lee, A. R. Kay, and D. W. Tank. Brain magnetic resonance imaging with
1411 contrast dependent on blood oxygenation. *Proceedings of the National Academy of
1412 Sciences*, 87(24):9868–9872, December 1990. ISSN 1091-6490. doi: 10.1073/pnas.
1413 87.24.9868. URL <http://dx.doi.org/10.1073/pnas.87.24.9868>.
- 1414
- 1415 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>.
- 1416
- 1417 Jonathan D. Power, Alexander L. Cohen, Steven M. Nelson, Gagan S. Wig, Kelly A.
1418 Barnes, Jessica A. Church, Alecia C. Vogel, Timothy O. Laumann, Fran M. Miezin,
1419 Bradley L. Schlaggar, and Steven E. Petersen. Functional Network Organization
1420 of the Human Brain. *Neuron*, 72(4):665–678, November 2011. ISSN 08966273.

- 1421 doi: 10.1016/j.neuron.2011.09.006. URL [http://dx.doi.org/10.1016/j.
1422 neuron.2011.09.006.](http://dx.doi.org/10.1016/j.neuron.2011.09.006)
- 1423 Jonathan D. Power, Kelly A. Barnes, Abraham Z. Snyder, Bradley L. Schlaggar, and
1424 Steven E. Petersen. Spurious but systematic correlations in functional connectivity
1425 MRI networks arise from subject motion. *NeuroImage*, 59(3):2142–2154, February
1426 2012. ISSN 1095-9572. doi: 10.1016/j.neuroimage.2011.10.018. URL [http://
1427 dx.doi.org/10.1016/j.neuroimage.2011.10.018](http://dx.doi.org/10.1016/j.neuroimage.2011.10.018).
- 1428 Jonathan D. Power, Anish Mitra, Timothy O. Laumann, Abraham Z. Snyder, Bradley L.
1429 Schlaggar, and Steven E. Petersen. Methods to detect, characterize, and remove mo-
1430 tion artifact in resting state fMRI. *NeuroImage*, 84:320–341, January 2014. doi:
1431 10.1016/j.neuroimage.2013.08.048. URL [http://dx.doi.org/10.1016/j.
1432 neuroimage.2013.08.048](http://dx.doi.org/10.1016/j.neuroimage.2013.08.048).
- 1433 Chris Rorden, Hans-Otto Karnath, and Leonardo Bonilha. Improving lesion-symptom
1434 mapping. *J Cogn Neurosci*, 19(7):1081–1088, Jul 2007. doi: 10.1162/jocn.2007.19.
1435 7.1081. URL <http://dx.doi.org/10.1162/jocn.2007.19.7.1081>.
- 1436 Ziad S. Saad, Stephen J. Gotts, Kevin Murphy, Gang Chen, Hang Joon J. Jo, Alex Martin,
1437 and Robert W. Cox. Trouble at rest: how correlation patterns and group differences
1438 become distorted after global signal regression. *Brain connectivity*, 2(1):25–32, 2012.
1439 ISSN 2158-0022. doi: 10.1089/brain.2012.0080. URL [http://dx.doi.org/
1440 10.1089/brain.2012.0080](http://dx.doi.org/10.1089/brain.2012.0080).
- 1441 R. Salvador, J. Suckling, M. R. Coleman, J. D. Pickard, D. Menon, and E. Bullmore.
1442 Neurophysiological architecture of functional magnetic resonance images of human
1443 brain. *Cereb Cortex*, 15(9):1332–1342, September 2005. ISSN 1047-3211. URL
1444 <http://view.ncbi.nlm.nih.gov/pubmed/15635061>.
- 1445 R. Salvador, A. Martínez, E. Pomarol-Clotet, S. Sarró, J. Suckling, and E. Bull-
1446 more. Frequency based mutual information measures between clusters of brain re-
1447 gions in functional magnetic resonance imaging. *NeuroImage*, 35(1):83–88, March

- 1448 2007. ISSN 1053-8119. doi: 10.1016/j.neuroimage.2006.12.001. URL <http://dx.doi.org/10.1016/j.neuroimage.2006.12.001>.
- 1450 Robert E. Schapire, Yoav Freund, Peter Bartlett, and Wee Sun Lee. Boosting the margin:
1451 a new explanation for the effectiveness of voting methods. *The Annals of Statistics*, 26
1452 (5):1651–1686, 10 1998. doi: 10.1214/aos/1024691352. URL <http://dx.doi.org/10.1214/aos/1024691352>.
- 1453
- 1454 Ganesh M Shankar, Shaomin Li, Tapan H Mehta, Amaya Garcia-Munoz, Nina E Shep-
1455 ardson, Imelda Smith, Francesca M Brett, Michael A Farrell, Michael J Rowan, Cyn-
1456 thia A Lemere, Ciaran M Regan, Dominic M Walsh, Bernardo L Sabatini, and Den-
1457 nis J Selkoe. Amyloid-beta protein dimers isolated directly from alzheimer's brains
1458 impair synaptic plasticity and memory. *Nat Med*, 14(8):837–842, Aug 2008. doi:
1459 10.1038/nm1782. URL <http://dx.doi.org/10.1038/nm1782>.
- 1460 Zarrar Shehzad, Clare M. Kelly, Philip T. Reiss, Dylan G. Gee, Kristin Gotimer, Lu-
1461 cina Q. Uddin, Sang Han H. Lee, Daniel S. Margulies, Amy Krain K. Roy, Bharat B.
1462 Biswal, Eva Petkova, F. Xavier Castellanos, and Michael P. Milham. The resting
1463 brain: unconstrained yet reliable. *Cerebral cortex (New York, N.Y. : 1991)*, 19(10):
1464 2209–2229, October 2009. ISSN 1460-2199. doi: 10.1093/cercor/bhn256. URL
1465 <http://dx.doi.org/10.1093/cercor/bhn256>.
- 1466 Yvette Sheline and Marcus Raichle. Resting state functional connectivity in pre-
1467 clinical alzheimerâŽs disease. *Biological Psychiatry*, 74(5):340–347, September
1468 2013. ISSN 00063223. URL <http://dx.doi.org/10.1016/j.biopsych.2012.11.028>.
- 1469
- 1470 Reisa A. Sperling, Paul S. Aisen, Laurel A. Beckett, David A. Bennett, Suzanne Craft,
1471 Anne M. Fagan, Takeshi Iwatsubo, Clifford R. Jack, Jeffrey Kaye, Thomas J. Mon-
1472 tine, Denise C. Park, Eric M. Reiman, Christopher C. Rowe, Eric Siemers, Yaakov
1473 Stern, Kristine Yaffe, Maria C. Carrillo, Bill Thies, Marcelle Morrison-Bogorad,
1474 Molly V. Wagster, and Creighton H. Phelps. Toward defining the preclinical stages

- 1475 of Alzheimer's disease: Recommendations from the National Institute on Aging-
1476 Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's dis-
1477 ease. *Alzheimer's & Dementia*, 7(3):280–292, May 2011. doi: 10.1016/j.jalz.2011.
1478 03.003. URL <http://dx.doi.org/10.1016/j.jalz.2011.03.003>.
- 1479 Lixia Tian, Tianzi Jiang, Meng Liang, Xiaobo Li, Yong He, Kun Wang, Bingli Cao, and
1480 Tao Jiang. Stabilities of negative correlations between blood oxygen level-dependent
1481 signals associated with sensory and motor cortices. *Hum Brain Mapp*, 28(7):681–
1482 690, Jul 2007. doi: 10.1002/hbm.20300. URL <http://dx.doi.org/10.1002/hbm.20300>.
- 1483
- 1484 Dardo Tomasi and Nora D. Volkow. Functional connectivity density mapping. *Pro-
1485 ceedings of the National Academy of Sciences*, 107(21):9885–9890, May 2010.
1486 doi: 10.1073/pnas.1001414107. URL <http://dx.doi.org/10.1073/pnas.1001414107>.
- 1487
- 1488 N. Tzourio-Mazoyer, B. Landeau, D. Papathanassiou, F. Crivello, O. Etard, N. Delcroix,
1489 B. Mazoyer, and M. Joliot. Automated anatomical labeling of activations in SPM
1490 using a macroscopic anatomical parcellation of the MNI MRI single-subject brain.
1491 *NeuroImage*, 15(1):273–289, January 2002. ISSN 1053-8119. doi: 10.1006/nimg.
1492 2001.0978. URL <http://dx.doi.org/10.1006/nimg.2001.0978>.
- 1493
- 1494 Martijn van den Heuvel, Rene Mandl, and Hilleke Hulshoff Pol. Normalized Cut
1495 Group Clustering of Resting-State fMRI Data. *PLoS ONE*, 3(4):e2001+, April 2008.
1496 doi: 10.1371/journal.pone.0002001. URL <http://dx.doi.org/10.1371/journal.pone.0002001>.
- 1497
- 1498 Martijn P. van den Heuvel, René C. Mandl, René S. Kahn, and Hilleke E. Hul-
1499 shoff Pol. Functionally linked resting-state networks reflect the underlying struc-
1500 tural connectivity architecture of the human brain. *Human brain mapping*, 30(10):
1501 3127–3141, October 2009. ISSN 1097-0193. doi: 10.1002/hbm.20737. URL
1502 <http://dx.doi.org/10.1002/hbm.20737>.

- 1502 Koene R. Van Dijk, Trey Hedden, Archana Venkataraman, Karleyton C. Evans, Sara W.
1503 Lazar, and Randy L. Buckner. Intrinsic functional connectivity as a tool for human
1504 connectomics: theory, properties, and optimization. *Journal of neurophysiology*, 103
1505 (1):297–321, January 2010. ISSN 1522-1598. doi: 10.1152/jn.00783.2009. URL
1506 <http://dx.doi.org/10.1152/jn.00783.2009>.
- 1507 G. Vanhoutte, M. Verhoye, and A. Van der Linden. Changing body temperature affects
1508 the t2* signal in the rat brain and reveals hypothalamic activity. *Magn Reson Med*,
1509 55(5):1006–1012, May 2006. doi: 10.1002/mrm.20861. URL <http://dx.doi.org/10.1002/mrm.20861>.
- 1511 Archana Venkataraman, Marek Kubicki, Carl-Fredrik Westin, and Polina Golland. Ro-
1512 bust feature selection in resting-state fmri connectivity based on population stud-
1513 ies. *Conf Comput Vis Pattern Recognit Workshops*, pages 63–70, 2010. doi: 10.
1514 1109/CVPRW.2010.5543446. URL [http://dx.doi.org/10.1109/CVPRW.](http://dx.doi.org/10.1109/CVPRW.2010.5543446)
1515 2010.5543446.
- 1516 Anthony B. Waites, Regula S. Briellmann, Michael M. Saling, David F. Abbott, and
1517 Graeme D. Jackson. Functional connectivity networks are disrupted in left temporal
1518 lobe epilepsy. *Annals of neurology*, 59(2):335–343, February 2006. doi: 10.1002/ana.
1519 20733. URL <http://dx.doi.org/10.1002/ana.20733>.
- 1520 Kun Wang, Meng Liang, Liang Wang, Lixia Tian, Xinqing Zhang, Kuncheng Li, and
1521 Tianzi Jiang. Altered functional connectivity in early Alzheimer’s disease: a resting-
1522 state fMRI study. *Human brain mapping*, 28(10):967–978, October 2007. ISSN
1523 1065-9471. doi: 10.1002/hbm.20324. URL <http://dx.doi.org/10.1002/hbm.20324>.
- 1525 Liang Wang, Yufeng Zang, Yong He, Meng Liang, Xinqing Zhang, Lixia Tian, Tao Wu,
1526 Tianzi Jiang, and Kuncheng Li. Changes in hippocampal connectivity in the early
1527 stages of Alzheimer’s disease: evidence from resting state fMRI. *NeuroImage*, 31
1528 (2):496–504, June 2006. ISSN 1053-8119. doi: 10.1016/j.neuroimage.2005.12.033.
1529 URL <http://dx.doi.org/10.1016/j.neuroimage.2005.12.033>.

- 1530 Cristen J. Willer, Yun Li, and Gonçalo R. Abecasis. METAL: fast and efficient
1531 meta-analysis of genomewide association scans. *Bioinformatics*, 26(17):2190–2191,
1532 September 2010. doi: 10.1093/bioinformatics/btq340. URL <http://dx.doi.org/10.1093/bioinformatics/btq340>.
- 1533
- 1534 A. M. Wink and J. B. T. M. Roerdink. BOLD noise assumptions in fMRI. *Int. J.
1535 Biomed. Imaging*, 2006, 2006. doi: 10.1155/IJBI/2006/12014. URL <http://dx.doi.org/10.1155/IJBI/2006/12014>.
- 1536
- 1537 K. J. Worsley and K. J. Friston. Analysis of fMRI Time-Series Revisited—Again. *NeuroImage*,
1538 2(3):173–181, September 1995. ISSN 10538119. doi: 10.1006/nimg.1995.
1539 1023. URL <http://dx.doi.org/10.1006/nimg.1995.1023>.
- 1540
- 1541 Chao-Gan Yan, Brian Cheung, Clare Kelly, Stan Colcombe, R. Cameron Craddock,
1542 Adriana Di Martino, Qingyang Li, Xi-Nian Zuo, F. Xavier Castellanos, and Michael P.
1543 Milham. A comprehensive assessment of regional variation in the impact of head
1544 micromovements on functional connectomics. *NeuroImage*, 76:183–201, August
1545 2013a. doi: 10.1016/j.neuroimage.2013.03.004. URL <http://dx.doi.org/10.1016/j.neuroimage.2013.03.004>.
- 1546
- 1547 Chao-Gan G. Yan, Cameron C. Craddock, Xi-Nian N. Zuo, Yu-Feng F. Zang, and
1548 Michael P. Milham. Standardizing the intrinsic brain: towards robust measurement
1549 of inter-individual variation in 1000 functional connectomes. *NeuroImage*, 80:246–
1550 262, October 2013b. URL <http://view.ncbi.nlm.nih.gov/pubmed/23631983>.
- 1551
- 1552 Chaogan Yan, Dongqiang Liu, Yong He, Qihong Zou, Chaozhe Zhu, Xinian Zuo, Xi-
1553 angyu Long, and Yufeng Zang. Spontaneous Brain Activity in the Default Mode Net-
1554 work Is Sensitive to Different Resting-State Conditions with Limited Cognitive Load.
1555 *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>.
- 1556 Hong Yang, Xiang-Yu Long, Yihong Yang, Hao Yan, Chao-Zhe Zhu, Xiang-Ping Zhou,

- 1557 Yu-Feng Zang, and Qi-Yong Gong. Amplitude of low frequency fluctuation within
1558 visual areas revealed by resting-state functional mri. *Neuroimage*, 36(1):144–152,
1559 May 2007. doi: 10.1016/j.neuroimage.2007.01.054. URL <http://dx.doi.org/10.1016/j.neuroimage.2007.01.054>.
- 1561 Zhiwen Yu, Hau-San Wong, and Hongqiang Wang. Graph-based consensus clustering
1562 for class discovery from gene expression data. *Bioinformatics*, 23(21):2888–2896,
1563 November 2007. doi: 10.1093/bioinformatics/btm463. URL <http://dx.doi.org/10.1093/bioinformatics/btm463>.
- 1565 Y. Yuan, Z-X. Gu, and W-S. Wei. Fluorodeoxyglucose-positron-emission tomography,
1566 single-photon emission tomography, and structural mr imaging for prediction of rapid
1567 conversion to alzheimer disease in patients with mild cognitive impairment: a meta-
1568 analysis. *AJNR Am J Neuroradiol*, 30(2):404–410, Feb 2009. doi: 10.3174/ajnr.A1357.
1569 URL <http://dx.doi.org/10.3174/ajnr.A1357>.
- 1570 Yu-Feng F. Zang, Yong He, Chao-Zhe Z. Zhu, Qing-Jiu J. Cao, Man-Qiu Q. Sui,
1571 Meng Liang, Li-Xia X. Tian, Tian-Zi Z. Jiang, and Yu-Feng F. Wang. Altered base-
1572 line brain activity in children with ADHD revealed by resting-state functional MRI.
1573 *Brain & development*, 29(2):83–91, March 2007. ISSN 0387-7604. doi: 10.1016/j.braindev.
1574 2006.07.002. URL <http://dx.doi.org/10.1016/j.braindev.2006.07.002>.
- 1576 H. Zhang, S. Wang, J. Xing, B. Liu, Z. Ma, M. Yang, Z. Zhang, and G. Teng. De-
1577 tector of PCC functional connectivity characteristics in resting-state fMRI in mild
1578 Alzheimer’s disease. *Behavioural Brain Research*, 197(1):103–108, January 2009.
1579 ISSN 01664328. doi: 10.1016/j.bbr.2008.08.012. URL <http://dx.doi.org/10.1016/j.bbr.2008.08.012>.
- 1581 Han Zhang, Xi-Nian Zuo, Shuang-Ye Ma, Yu-Feng Zang, Michael P. Milham, and
1582 Chao-Zhe Zhu. Subject Order-Independent Group ICA (SOI-GICA) for Func-
1583 tional MRI Data Analysis. *NeuroImage*, March 2010. ISSN 10538119. doi:

- 1584 10.1016/j.neuroimage.2010.03.039. URL [http://dx.doi.org/10.1016/j.
1585 neuroimage.2010.03.039.](http://dx.doi.org/10.1016/j.neuroimage.2010.03.039)
- 1586 Yuan Zhong, Huinan Wang, Guangming Lu, Zhiqiang Zhang, Qing Jiao, and Yi-
1587 jun Liu. Detecting Functional Connectivity in fMRI Using PCA and Regres-
1588 sion Analysis. *Brain Topography*, 22(2):134–144, September 2009. ISSN 0896-
1589 0267. doi: 10.1007/s10548-009-0095-4. URL [http://dx.doi.org/10.
1590 1007/s10548-009-0095-4](http://dx.doi.org/10.1007/s10548-009-0095-4).
- 1591 Yuan Zhou, Meng Liang, Lixia Tian, Kun Wang, Yihui Hao, Haihong Liu, Zhenling Liu,
1592 and Tianzi Jiang. Functional disintegration in paranoid schizophrenia using resting-
1593 state fmri. *Schizophr Res*, 97(1-3):194–205, Dec 2007. doi: 10.1016/j.schres.2007.
1594 05.029. URL <http://dx.doi.org/10.1016/j.schres.2007.05.029>.
- 1595 Yuan Zhou, Ni Shu, Yong Liu, Ming Song, Yihui Hao, Haihong Liu, Chunshui Yu, Zhen-
1596 ing Liu, and Tianzi Jiang. Altered resting-state functional connectivity and anatomi-
1597 cal connectivity of hippocampus in schizophrenia. *Schizophr Res*, 100(1-3):120–132,
1598 Mar 2008. doi: 10.1016/j.schres.2007.11.039. URL [http://dx.doi.org/10.
1599 1016/j.schres.2007.11.039](http://dx.doi.org/10.1016/j.schres.2007.11.039).
- 1600 A. P. Zijdenbos, R. Forghani, and A. C. Evans. Automatic "pipeline" analysis of
1601 3-D MRI data for clinical trials: application to multiple sclerosis. *IEEE Trans-
1602 actions on Medical Imaging*, 21(10):1280–1291, October 2002. ISSN 0278-0062.
1603 doi: 10.1109/TMI.2002.806283. URL [http://dx.doi.org/10.1109/TMI.
1604 2002.806283](http://dx.doi.org/10.1109/TMI.2002.806283).
- 1605 Xi-Nian Zuo, Ting Xu, Lili Jiang, Zhi Yang, Xiao-Yan Cao, Yong He, Yu-Feng Zang,
1606 F. Xavier Castellanos, and Michael P. Milham. Toward reliable characterization of
1607 functional homogeneity in the human brain: Preprocessing, scan duration, imaging
1608 resolution and computational space. *NeuroImage*, October 2012. ISSN 10538119. doi:
1609 10.1016/j.neuroimage.2012.10.017. URL [http://dx.doi.org/10.1016/j.
1610 neuroimage.2012.10.017](http://dx.doi.org/10.1016/j.neuroimage.2012.10.017).