UNIVERSITÀ DEGLI STUDI
DI TRENTO

CIMeC - Center for Mind/Brain Sciences

PhD Dissertation

Test-retest reliability of intrinsic human brain default-mode fMRI connectivity: a study of slice acquisition and physiological noise correction effects

Author:

Rocco MARCHITELLI

Advisor:

Dr. Jorge JOVICICH

A THESIS SUBMITTED FOR THE DEGREE OF

PHILOSOPHIAE DOCTOR (PHD)

DOCTORAL SCHOOL IN COGNITIVE AND BRAIN SCIENCES

XXVIII CYCLE

April, 2016

To Maris

with kind affection

Contents

Declaration

I herewith declare that I have produced this paper without the prohibited assistance of third parties and without making use of aids other than those specified; notions taken over directly or indirectly from other sources have been identified as such. This thesis has not previously been presented in identical or similar form to any other Italian or foreign examination board. The thesis work was conducted under the supervision of Jorge Jovicich at the University of Trento.

Trento, February 2016

Acknowledgements

I would like to thank my parents for the longstanding support provided during these years and all the people who guided me through this important path and shared this experience with me. It was very important for me to find professional researchers who invested their expertise in me with passion and steadfastness, believing in my dedication to cognitive neuroimaging.

I express unlimited gratitude to my advisor, Dr. Jorge Jovicich, and the members of the oversight committee, Dr. Uri Hasson and Dr. Olivier Collignon, for methodically supervising my doctoral works. I herewith also thank all the other doctors of the MRI method group, Domenico Zaca, Ludovico Minati, Paolo Ferrari and the LNIF staff for providing precious help when needed. Going back in time, I am bound to recognize the contribution of Ben Davis who introduced me to fMRI data analysis, de facto shaping the way to my doctoral studies.

I am grateful to all members and collaborators of the PharmaCog project, funded by the EU-FP7 for the Innovative Medicine Initiative (grant no. 115009). Moreover, special thanks go to Giulia Elli for resting-state data acquisition and to the CiMEC MRI laboratory for medical and technical support. Research was supported in part by a European Research Council starting grant (MADVIS; ERC-StG 337573) attributed to O. Collignon.

I give thanks to all the doctoral colleagues and students whom I have had the plasure to spend my time with. This fascinating international environment allowed me to get closer to different people and cultures from all over the world and discover their own characteristics and peculiarities. I take this opportunity to recall all my flatmates Eduard, Ulisse, Alberto, Reshanne, Evelyn, Moahmed, Marilina, Caterina, Matthias and all my friends Mauro, Yagmur, Dario, Betul, Yanina, Lorilei, Delyan, Katya, Anuch, Kusala, Megan, Kristen, Ronny, Klarissa, Deniz, Vania and Adrià. Forgive me for missing somebody and remember that each of you has contributed to my personal growth and to forge my future.

1

Introduction

1.1. Motivations & Goals

1.1.1. General summary

Resting-state or task-free functional magnetic resonance imaging (RS-fMRI) of the human brain is a functional neuroimaging technique that exploits the magnetic resonance (MR) technology to measure spontaneous brain activity in-vivo using endogenous contrasts associated with blood flow changes (Ogawa et al., 1990). fMRI experiments are characterized by the acquisition of T1-weighted (structural) and T2*-weighted (functional) images, the latter obtained using the blood-oxygen-level-dependent (BOLD) contrast (Ogawa et al., 1992). This endogenous contrast is advantageous for scientific and medical research for being non-invasive and comparatively cheap for not requiring administration of paramagnetic external contrasts or radioactive tracers (Ogawa et al., 1990).

The RS-fMRI methodology was introduced a couple of decades ago with the discovery of functional co-activation patterns in sensory-motor areas even in absence of task performance (Biswal et al., 1995; Biswal et al., 1997) and has since been developed into well structured experimental protocols. During RS-fMRI studies, subjects are typically instructed to stay still in the scanner, refraining from being engaged in any cognitive task. This allows the investigation of the whole human brain intrinsic functional organization across several neurophysiological states.

This thesis will consider functional connectivity MRI (FC-fMRI) measures of cortical brain activity using RS-fMRI. FC-fMRI is defined as a statistical measure of temporal dependencies of the BOLD signal (defined on T2*-weighted images) among distinct brain regions or areas of the brain (defined on T1-weighted images). Resting-state FC-fMRI revealed the existence of several widely distributed coactive brain areas,

usually referred to as resting-state or intrinsic brain networks that are associated with self-oriented cognition and other neurophysiologic and metabolic processes in the cortical areas under consideration (Van den Heuvel & Hulshoff Pol, 2010).

In this thesis we focus on one particular intrinsic brain network, the default-mode network (DMN) (Greicius, 2003). The DMN has recently received particular attention from clinical neurosciences since its intrinsic FC-fMRI appears to be sensitive to a wide range of neurodevelopmental, neuropsychiatric and neurodegerative disorders (Anticevic et al., 2012). Therefore, considering the ease of acquisition and the technical feasibility of RS-fMRI with non-cooperative clinical populations, intrinsic DMN connectivity properties could be used to validate clinical and develop preclinical biomarkers to predict and monitor disease progression (Chhatwal & Sperling, 2012; Greicius et al., 2004; Zhou et al., 2010).

Despite the potentials of the intrinsic DMN connectivity as a disease marker, there are several signal fluctuations of non-neural origins that confound intrinsic FC-fMRI estimates within the DMN. The BOLD contrast is only an indirect measure of cortical brain activity and captures several unwanted artifacts rising from human physiology or MRI hardware system. This means that suboptimal choices in defining RS-fMRI protocols could lead to reduced BOLD sensitivity to neural activity. Recently, many issues related to different aspects of RS-fMRI protocol were investigated, including acquisition issues (Biswal et al., 2010; Patriat et al., 2013), data preprocessing (Cordes et al., 2001) and analysis (Chang et al., 2009; Fox et al., 2009; Murphy et al., 2009; Power et al., 2012; Tzourio-Mazoyer et al., 2002; Van Dijk et al., 2012; Yeo et al., 2011). Altogether, these observations indicate that physiological and MRI hardware related artifacts confound FC-fMRI metrics in RS-fMRI making their interpretation quite challenging (Power et al., 2014; Van Dijk et al., 2012; Weissenbacher et al., 2009). This thesis focuses on the identification and elimination of two main resting-state FC-fMRI confounds that hamper the characterization of intrinsic FC-fMRI within the default-mode system: physiological noise (intended as cardiac and respiratory related fluctuations) and head motion.

1.1.2. Physiological noise artifacts

Physiological noise rising from cardiac activity and respiration can either increase or decrease intrinsic FC-fMRI in resting-state networks leading to an overestimation or underestimation of intrinisc DMN connectivity (Churchill et al., 2012; Liu et al., 2006; Welvaert & Rosseel, 2012). In particular, unless monitored during the fMRI scanning sessions and removed offline (Birn et al., 2006; Birn et al., 2009; Lund, 2001) signals associated with human physiology unavoidably introduce unwanted variability in resting-state FC-fMRI metrics across subjects or repeated sessions. However, due to technical limitations or lack of MRI equipment it is often not possible to monitor physiological signals during MR acquisitions. This is often the case in multisite studies where MRI-compatible pulse-oxymeters and respiration belts might not be available at each site.

In these cases, retrospective data-based correction methods can be used to estimate and reduce the effects of physiological noise. These technical approaches include the averaged white matter (WM) and cerebral spinal fluid (CSF) nuisance regressions (Behzadi et al., 2007; Jo et al., 2010), Bayesian methods to track the frequency trajectories of cardiac and respiration for removing physiological noise (Sarkka et al., 2012) or canonical correlations analysis to identify autocorrelated physiological noise (Churchill & Strother, 2013). Other methods exploit independent component analysis (ICA) to remove temporal (Beall, 2010; Beall & Lowe, 2007) or spatial components associated to physiological noise (Griffanti et al., 2014; Salimi-Khorshidi et al., 2014).

Retrospective physiological denoising methods have their limitations. Retrospective estimation of physiological fluctuations based on the periodicity of the respiration or the heart beat (Glover et al., 2000; Hu et al., 1995) is limited by the model-order selection of the Fourier transformations (Beall & Lowe, 2007). While low model-orders would only minimize noise in the data (Glover et al., 2000) higher model-orders would overfit heart beat and respiration profiles (Harvey et al., 2008) reducing also non-noise signal variance (Beall & Lowe, 2007). On the other hand, ICA-based separation does not reduce temporal correlations between resting-state brain networks and parallel measures of human physiology (Beall & Lowe, 2010). This could depend

on ICA model-order dimensionality (Beall & Lowe, 2010) and highlights the need of a specific physiological denoising strategy prior running ICA algorithms.

Therefore, it remains an open question what is an optimal physiological correction strategy that increases DMN specificity. The contributions of cardiac and respiratory activity on BOLD signal fluctuations as a function of age are unknown; therefore physiological effects in young and elderly people may be different. Furthermore, the application of resting-state FC-fMRI as a longitudinal biomarker of dementia would require both reliability and sensitivity to longitudinal neuropathological changes. An optimal physiological noise correction is expected to improve the stability of DMN connectivity measures over time, thereby increasing the sensitivity of longitudinal studies evaluating DMN connectivity changes related to normal aging, disease progression or disease treatment.

1.1.3. Head-motion artifacts

Another important confounding factor in RS-fMRI comes from head movements, even if small, which can introduce spurious but structured noise in functional brain connectivity measures (Power et al., 2015). Movements typically induce reduction and increase in FC-fMRI metrics between spatially distant and close areas, respectively. Considering the main regions of the DMN, connectivity between anterior-posterior regions is particularly vulnerable to head motion (Power et al., 2012; Van Dijk et al., 2012). Connectivity loss induced by head-motion effects in the anterior-posterior regions of the DMN could lead to erroneous diagnosis of pathological conditions characterized by reduction of connectivity in the anterior regions of the network. In fact, motion effects are particularly problematic in children (Fair et al., 2007) and elderly (Andrews-Hanna et al., 2007) individuals who are generally characterized by higher motion than healthy young adults. This limits motion-unbiased cross-sectional comparisons of intrinsic DMN in development and aging.

Furthermore, longitudinal evaluations demonstrated a predisposition to motion in young adults (Van Dijk et al., 2012). This indicates that motion is not exclusively incidental but more broadly constitutes an individual biological trait (Van Dijk et al., 2012; Zeng et al., 2014). It remains challenging to distinguish between “biological” and “incidental” head-motion because no “gold standard” is available to date.

There are two general types of approaches aimed at addressing head motion correction in MRI studies, one is based on prospective “real-time” measures of head motion with reorientation of the imaging slices during acquisition, and the other is based on retrospective estimation and reorientation of the data during processing of the data after acquisition (Godenschweger et al., 2016). The study of prospective measures is a promising active field of research, but such methods are not yet standardly available. Here we focus on retrospective head motion correction methods, moreover in some aspects of data acquisition.

Standard retrospective motion correction approaches implemented in FC-fMRI studies account only for rigid head-motion effects occurring between the acquired volumes (Jenkinson et al., 2002). However, in standard single-shot 2D echo-planar imaging (EPI) acquisition protocols there are two acquisition choices, interleaved or sequential acquisitions. One open question is whether there are quantitative advantages between using one of these methods for motion sensitivity effects on RS-fMRI, in particular it is not clear how slice-timing differences between the acquisition of adjacent slices might influence BOLD sensitivity to head-motion and consequently intrinsic DMN connectivity.

A separate issue relates to the implementation of the co-registrations done for the head motion correction, which can be performed on the full 3D volume or at the 2D slice level. Standard volumetric correction methods are therefore not sensitive to true motion occurring during the acquisition of each single volume. Although novel methods have been recently introduced to correct for deformable (non-rigid) motion within each acquired volume (Beall & Lowe, 2014) more effort should be made to account for through-plane motion effects (Godenschweger et al., 2016). These technical limitations are potentially relevant in longitudinal study designs since motion might be associated with reduced FC-fMRI reliability in the DMN (Guo et al., 2012).

1.1.4. Thesis goals

This thesis aims at evaluating, in two separate studies, strategies for physiological noise and head motion correction in resting state brain FC-fMRI. In particular, as a general marker of noise correction performance we use the test-retest reproducibility of the DMN. The guiding hypothesis is that methods that improve

reproducibility should reflect more efficient corrections and thus be preferable in longitudinal studies.

The physiological denoising study evaluated longitudinal changes in a 3T harmonized multisite fMRI study of healthy elderly participants from the PharmaCog Consortium (Jovicich et al., 2016). Retrospective physiological noise correction (rPNC) methods were here implemented to investigate their influence on several DMN reliability measures within and between 13 MRI sites. Each site involved five different healthy elderly participants who were scanned twice at least a week apart (5 participants per site). fMRI data analysis was performed once without rPNC and then with WM/CSF regression, with physiological estimation by temporal ICA (PESTICA) (Beall & Lowe, 2007) and FMRIB's ICA-based Xnoiseifier (FSL-FIX) (Griffanti et al., 2014; Salimi-Khorshidi et al., 2014). These methods differ for their data-based computational approach to identify physiological noise fluctuations and need to be applied at different stages of data preprocessing. As a working hypothesis, physiological denoising was in general expected to improve DMN reliability.

The head motion study evaluated longitudinal changes in the DMN connectivity from a 4T single-site study of 24 healthy young volunteers who were scanned twice within a week. Within each scanning session, RS-fMRI scans were acquired once using interleaved and then sequential slice-order acquisition methods. Furthermore, brain volumes were corrected for motion using once rigid-body volumetric and then slice-wise methods. The effects of these choices were then evaluated computing multiple DMN reliability measures and investigating single regions within the DMN to assess the existence of inter-regional effects associated with head-motion. In this case, we expected to find slice-order acquisition effects in reliability estimates under standard volumetric motion correction and no slice-order acquisition effect under 2D slice-based motion correction.

Both studies used ICA to characterize the DMN using group-ICA and dual regression procedures (Beckmann et al., 2009). This methodology proved successful at defining consistent DMN connectivity metrics in longitudinal and clinical RS-fMRI studies (Zuo & Xing, 2014). Automatic DMN selection procedures and other quality assurance analyses were made to supervise ICA performance.

Both studies considered several test-retest (TRT) reliability estimates (Vilagut, 2014) for some DMN connectivity measurements: absolute percent error between the sessions, intraclass correlation coefficients (ICC) between sessions and multiple sites, the Jaccard index to evaluate the degree of voxel-wise spatial pattern actiavtion overlap between sessions.

1.2. Research Outline

After a general introduction (Chapter 2), the research in this thesis focuses on the evaluation of physiological noise correction methods (Chapter 3) and on the evaluation of acquisition and analyses strategies to reduce the effects of head motion (Chapter 4).

The physiological noise study (Chapter 3) revealed that retrospective physiological denoising methods significantly affected the mean z-scores and, albeit less markedly, the cluster-size in the DMN; in particular, FSL-FIX tended to increase the DMN z-scores compared to others. Within-site test-retest reliability was consistent across sites, with no differences across denoising methods. The absolute percent errors were in the range of 5-11% for DMN z-scores and cluster-size reliability. DMN pattern overlap was in the range 60-65%. In particular, no rPNC method showed a significant reliability improvement compared with no physiological correction implemented. However, FSL-FIX and WM/CSF regressions showed both similar and significant improvements of reproducibility consistency across the consortium (ICC = 0.67) for the DMN z-scores relative to no physiological noise correction (NPC). Overall these findings support the use of rPNC methods like WM/CSF regressions or FSL-FIX to characterize multisite longitudinal changes of intrinsic FC-fMRI.

The head motion study (Chapter 4) revealed that mean z-scores within the DMN are influenced by both slice-order acquisition and motion correction methods in all DMN regions even in presence of low motion. In contrast, no combinations-of-interest influenced or systematically improved the TRT reliability in all regions. In the DMN, TRT reliability errors were overall below 8% and ICC were overall moderate 0.47 (C.I. 0.29-0.57), indicating longitudinally stable spatio-temporal network characteristics. The average DMN pattern overlap was 40% (range: 14-65%). The longitudinal spatial reproducibility was the lowest in the ACC region at 30%. Acquisition protocols and motion correction methods could however be optimized to reduce cluster-size

variability in frontal DMN areas. These results support freedom of choice between the examined acquisition protocols and head-motion correction methods in longitudinal DMN studies.

The main thesis findings are outlined and put in perspective with the literature in the discussion (Chapter 5). These indicate that the intrinsic DMN connectivity is significantly sensitive to the factorts manipulated (slice-order acquisition and preprocessing correction choices), while TRT reliability of the DMN connectivity is not, or at least not systematically.

1.3. Contributions

The entire work presented in this thesis has been realized in the Laboratory for the Functional Neuroimaging (Lnif) within the Interdepartmental Center for Brain/Mind Sciences (CIMeC, University of Trento). Within the duration of the PhD activity, contributions were given with extensive data analysis to two main projects, namely PharmaCog, a European multisite consortium of Alzheimer research, and a resting-state Lnif project. The work performed during the PhD resulted in published works as well as several proceedings, in national and international conferences, both as first author or co-author. A list of author's publications is provided in the following.

1.3.1. Publications

Jovicich J, Minati L, Marizzoni M, **Marchitelli R**, Sala-Llonch R, Bartrés-Faz D, Arnold J, Benninghoff J, Fiedler U, Roccatagliata L, Picco A, Nobili F, Blin O, Bombois S, Lopes R, Bordet R, Sein J, Ranjeva JP, Didic M, Gros-Dagnac H, Payoux P, Zoccatelli G, Alessandrini F, Beltramello A, Bargalló N, Ferretti A, Caulo M, Aiello M, Cavaliere C, Soricelli A, Parnetti L, Tarducci R, Floridi P, Tsolaki M, Constantinidis M, Drevelegas A, Rossini PM, Marra C, Schönknecht P, Hensch T, Hoffmann KT, Kuijer JP, Visser PJ, Barkhof F, Frisoni GB; PharmaCog Consortium. *Longitudinal reproducibility of default-mode network connectivity in healthy elderly participants: A multicentric resting-state fMRI study*. Neuroimage. 2016 Jan 1;124(Pt A):442-54.

Marchitelli R, Minati L, Marizzoni M, Bosch B, Bartrés-Faz D, Müller BW, Wiltfang J, Fiedler U, Roccatagliata L, Picco A, Nobili F, Blin O, Bombois S, Lopes R, Bordet R, Sein J, Ranjeva JP, Didic M, Gros-Dagnac H, Payoux P, Zoccatelli G, Alessandrini F, Beltramello A, Bargalló N, Ferretti A, Caulo M, Aiello M, Cavaliere C, Soricelli A, Parnetti L, Tarducci R, Floridi P, Tsolaki M, Constantinidis M, Drevelegas A, Rossini PM, Marra C, Schönknecht P, Hensch T, Hoffmann KT, Kuijer JP, Visser PJ, Barkhof F, Frisoni GB, Jovicich J; The Pharmacog Consortium. *Test-Retest Reliability of the Default Mode Network in a Multi-centric fMRI Study of Healthy Elderly: Effects of Data-Driven Physiological Noise Correction Techniques*, Human Brain Mapping, March 2016. doi: 10.1002/hbm.23157.

Marchitelli R, Collignon O. M. C., Jovicich J; *Test-Retest Reproducibility of the intrinsic Default Mode Network: Influence of Head-Motion Correction Methods and fMRI Slice-Order Acquisition Effects*. (In preparation).

1.3.2. Conferences and Workshops

Marchitelli R, Jovicich J, Marizzoni M, Bosch B, Bartrés-Faz D, Bargalló N, Roccatagliata L, Picco A, Nobili F, Gross-Dagnac H, Payoux P, Frisoni G, *Physiological Noise Correction Effects on Multisite Functional Connectivity Reproducibility*; Organization for Human Brain Mapping, Hamburg, Germany, June 8-12, 2014. (Poster presentation.)

Jovicich J, Minati L, **Marchitelli R**, Frisoni G, The Pharmacog Consortium, *Resting state functional connectivity in the default mode network: preliminary evaluation of multicenter test-retest reproducibility*, Biennial Conference on Resting State/Brain Connectivity, Boston, U.S.A, September 11-13, 2014. (Poster presentation.)

Jovicich J, **Marchitelli R**, Marizzoni M, Sala-Llonch R, Bartrés-Faz D, Bargalló N, Frisoni G, *Influences of physiological noise correction on the functional connectivity of the default mode network and its reproducibility using TC-GICA*, 13th International Geneva Springfield Symposium on Advances in Alzheimer Therapy, Geneva,

Switzerland March 26-29, 2014. (Poster presentation.)

Marchitelli R, Collignon O, Jovicich J. *Influence of Head Motion and fMRI Slice Acquisition on Intrinsic Default Mode Network Reliability*. Organization for Human Brain Mapping, Honolulu, Hawaii, USA, June 14-18, 2015. (Poster presentation.)

Marchitelli R, *Influence of data-driven physiological correction methods on the longitudinal reliability and multicentric consistency of the default mode network in healthy aging*. The 6th Italian Chapter of the International Society for Magnetic Resonance in Medicine, Verona, Italy, April 16-17, 2015. (Oral presentation.)

2

Resting-State Functional Connectivity MRI

2.1. Discovery and properties of resting-state networks

In their pioneeristic work, Biswal et al. were the first who observed spontaneous BOLD oscillations in the motor system during resting conditions marking de facto the beginning of RS-fMRI (Biswal et al., 1995). Later, further investigations confirmed the relevance of RS-fMRI as a valuable paradigm to investigate brain function (Biswal et al., 1997; Lowe et al., 1998) and functional mapping was extended to other areas throughout the whole human brain (Xiong et al., 1999).

These studies attempted to functionally map the entire brain using the intrinsic FC-MRI methodology. Intrinsic FC-fMRI is defined as a statistical measure of temporal dependencies between low-frequency BOLD signal oscillations from distinct brain areas, that is commonly obtained using temporal correlations (**Figure 2.1**) (Biswal et al., 2010; Fox & Raichle, 2007; Lowe et al., 2000). Intrinsic FC-fMRI was determinant in the exploration of the whole-brain intrinsic architecture and led to the discovery of many resting-state networks (**Figure 2.2**). These networks show resilient spatio-temporal characteristics which allows their robust identification across individuals (Damoiseaux et al., 2006; De Luca et al., 2006; Salvador et al., 2005). Some overlap with sensory-motor activations detected in task-based fMRI studies (Lee et al., 2013) which indicates that spontaneous brain activity during rest is relevant to understand cognition (Buckner & Vincent, 2007).

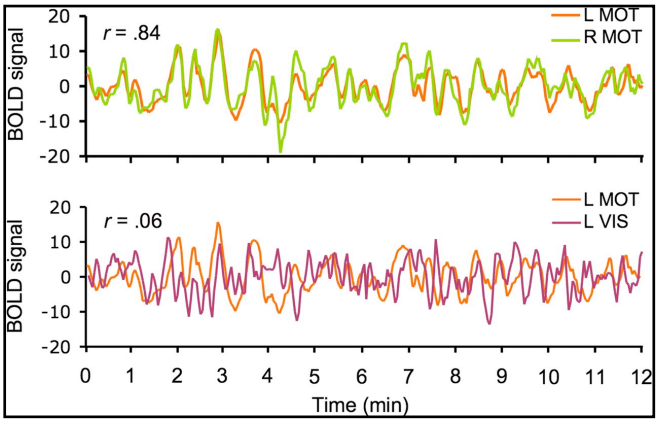

Figure 2.1. Example of intrinsic FC-fMRI. The figure shows temporal correlations of the low-frequency BOLD oscillations between three different regions. These were defined in the Left, L and Right, R Motor cortex, MOT and in the Left, L Visual cortex,VIS. High correlation between the inter-hemispheric motor corteces (top) and no correlation between ipsilateral correlations in left motor and visual cortex (bottom). from Van Dijk et al., 2010.

Resting-state networks are characterized by complex and dynamic interactivity between the regions involved in the network as well as with other networks (Kelly, 2008). These interactions are however difficult to be understood using only intrinsic FC-fMRI methods both because correlations are not very informative about the directionality of interactions within and between intrinsic brain systems, and because non-neural fMRI signals or some effects of particular data analysis methods can induce temporal correlations between two given time courses (Liao, 2009; Murphy et al., 2009). These limitations will be outlined in more details in the next sections since they represent the main concern of our experimental studies aimed at clarifing how to minimize the impact of non-neural spurious correlations from intrinsic FC-fMRI metrics.

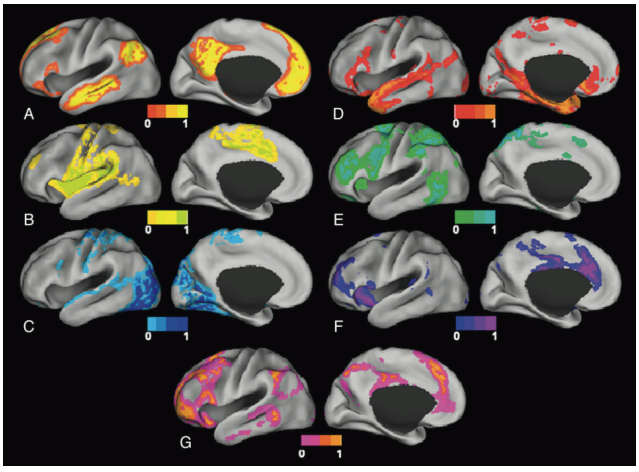

Figure 2.2. Consistent resting-state networks in the human brain. The figure shows the main widely distributed brain networks identified using RS-fMRI and projected on a standardized brain surface. These networks are the A) default-mode network; B) somatomotor network; C) visual network; D) language network; E) dorsal attention network; F) ventral attention network; G) frontoparietal network. From: (Lee et al., 2013)

2.2. Origins of resting-state fMRI signals

To date, the neural underpinnings of low-frequency oscillations in the RS-fMRI and consequently the neural sensitivity of intrinsic FC-fMRI mapping are not fully understood (Horwitz et al., 2005; Van den Heuvel & Hulshoff Pol, 2010). As previously mentioned, the fMRI signal is not a direct measure of neural activity but rather a measure of blood oxygenation, therefore BOLD effects do not only raise from neural activity, primary linked to synaptic metabolic activity, but also from random nuisance factors such as thermal or quantum mechanical noise and from structured nuisance components associated with signal reconstruction and distortion, non-neural physiological processes (i.e. cardiac and respiratory activity) and head-motion.

According to these considerations, the characteristics of resting-state and task-based fMRI signals would be alike with both advantages and disadvantages associated with the implementation of the BOLD contrast. However, since an evoked stimulus is missing in the resting-state, it is even more challenging to interpret and quantify the contributions of each constituent to the resulting BOLD effect (Arthurs & Boniface, 2002). For this reason, whether spontaneous oscillations include a neural constituent has been debated for long. Some have argued that they do not reflect neural activity but only physiological processes related to cardiac and respiratory activity (Birn, 2012; Birn et al., 2008; Shmueli et al., 2007). In contrast, some others have not denied the influence of physiological aliasing on intrinsic FC-fMRI but argued that spontaneous brain activity would also be neural in origin (Buckner & Vincent, 2007; Greicius et al., 2003; Gusnard et al., 2001) since these spontaneous oscillations mainly occur at very low frequencies (< 0.1 Hz) whereas physiological processes occur at higher frequencies (> 0.3 Hz) (Cordes et al., 2001; Cordes et al., 2000).

Several other reasons in favor of a neural hypothesis for the RS-fMRI signals have been proposed. Some were already outlined in the previous sections of this work and include the consistent identification of resting-state networks across healthy individuals (Damoiseaux et al., 2006) as well as the spatial correspondence with task-evoked networks in motor, visual and auditory areas which suggests that the intrinsic functional architecture of the brain develops in regions more frequently recruited for cognitive performance (Veer et al., 2010).

Finally, other reasons include the inter-hemispheric synchrony found in humans and other mammals (Biswal & Kannurpatti, 2009) and additional evidence comes from other imaging modalities such as electroencephalography (EEG): local field potentials at low-frequency also show temporal synchronicity (Kenet et al., 2003; Leopold & Logothetis, 2003; Lowe, 2012), recordings of neuronal firing are indirectly associated to the amplitude profiles of RS-fMRI correlations (Nir et al., 2008) and simultaneously measured fluctuations in neuronal spiking are associated to spontaneous BOLD fluctuations (Pan et al., 2013; Shmuel & Leopold, 2008; Shmuel et al., 2002).

In summary, both neural and non-neural fMRI signals contribute to intrinsic FC-fMRI. This highlights the need for optimizing the RS-fMRI design, developing technical solutions or improving analytical methods to minimize the influence of non-neural BOLD signals from intrinsic FC-fMRI measures.

2.3. The Default Mode Network

Among all resting-state functional networks mentioned previously, the work outlined in this thesis focuses on an intrinsic network in particular: the default mode network (DMN). Functional coactivations associated with this network typically encompass regions in the posterior cingulate cortex (PCC) and precuneus, medial prefrontal cortex (MPFC), anterior cingulate cortex (ACC), inferior parietal lobule and bilateral parietal cortex (LPC) (Buckner, 2012) (**Figure 2.3**). Neuroimaging research of the DMN grows at fast pace today and is grounded on solid scientific knowledge about its intrinsic functional dynamics and relevance for cognition (Buckner, 2012).

The DMN was observed for the first time in positron emission tomography (PET) studies (Mazoyer et al., 2001; Shulman et al., 1997) and the idea of a “default-mode” system was not immediately recognized until evidence in support of a plausible self-referential processing role was put forward (Gusnard & Raichle, 2001) and knowledge about its complex physiological dynamics was also gained from other imaging modalities (Laufs et al., 2003). Afterwards, first descriptions of the DMN using RS-fMRI were made using intrinsic FC-fMRI methods similar to those Biswal had used for mapping the motor system at rest in 1995 (Greicius, 2003).

Figure 2.3. Main regions of the DMN. Current representation of the four main regions of the DMN located in the posterior cingulate, bilateral temporo-parietal cortex and medial prefrontal cortex/anterior cingulate (Photo/Courtesy of Sarah Gimbel/USC).

The DMN owes its name to its main characteristic of showing prominent activation during the resting-state and low-demanding passive stimulation tasks in contrast to goal-directed cognitive tasks (Greicius & Menon, 2004). The DMN is a unique intrinsic brain network with this endogeneous dynamical dichotomy and for this reason it has been studied extensively. These DMN dynamics are crucial to understand its role in cognition: since the beginning, intrinsic FC-fMRI of the DMN was linked to memory function (Andreasen et al., 1995; Binder et al., 1999) and in particular to the retrieval of information from long-term memory, conscious representation of mental imagery and thoughts (Andrews-Hanna, 2012) and cognitive operations involved in problem solving and future planning (Buckner et al., 2008; Buckner & Carroll, 2007). In fact, it has been shown that the exacerbated dichotomy in DMN dynamics between attention-demanding tasks and resting-state (Fox et al., 2005) is mitigated by the inclusion of task features related to self-referential thoughts (McGuire et al., 1996), episodic memory (Shannon & Buckner, 2004) and meditation practice (Brewer et al., 2011; Jao et al., 2015).

In general, evidence suggests that the DMN is not innate. There are no adult-like patterns resembling the DMN in neonates and children aged below 4 years. Anyway, this might result from absent distant connectivity in children, which does not exclude the hypothesis of independent default mode function confined to single brain regions (Fair et al., 2009). Additional evidence suggests that DMN activity tends to decrease as a function of normal aging with substantial loss of intrinsic FC-fMRI (Andrews-Hanna et al., 2007; Damoiseaux et al., 2008; Ferreira et al., 2015; Sambataro et al., 2010). The most credited hypothesis is that loss of intrinsic FC-fMRI might be associated to age-related widespread structural changes (Horn et al., 2014) that affect the network efficiency in the transmission of neural information (Marstaller et al., 2015). Finally,

there is evidence suggesting that DMN connectivity is sensitive to circadian rhythm gradually reducing co-activations from morning to afternoon (Blautzik et al., 2013; Hodkinson et al., 2014).

2.4. Current issues in intrinsic DMN connectivity studies

RS-fMRI data analysis is not subjected to standardized acquisition and analysis methods. This lessens the reproducibility of findings across studies and increases the risk of adopting suboptimal acquisition and analysis methods (Griffanti et al., 2016) which could bias DMN connectivity metrics. On the other side the optimization of RS-fMRI data acquisition and analysis is challenging. It is not trivial to identify analytic expedients which favor the characterization of the DMN and the interpretability of its dynamics because of many issues arising from limitations in the fMRI technology and analysis software. The most critical ones are outlined in the following.

2.4.1. Data acquisition factors that affect DMN connectivity

Primarily, data acquisition choices can influence the overall data quality and determine decision making for analytic operations to be performed downstream including data preprocessing and FC-fMRI method choices. Standard acquisition protocols include single-shot multislice echo-planar imaging (EPI) which allows overall moderate temporal resolution (~ 0.5 Hz) in acquired volumes. This privileges the spatial dimension to perform data analysis and to compute FC-fMRI, unavoidably increasing the risk of spatial overlapping between the DMN and other brain networks or, even worse, structured noise (Birn et al., 2006). Novel acquisition sequences such as multiband echo-planar imaging allow the acquisition of multiple slices at the same time, improving the overal temporal resolution of fMRI (~ 1.25 Hz). This acquisition method improves DMN sensitivity (Feinberg et al., 2010) and has been used to investigate temporal modes of the DMN (Smith et al., 2012) but it is limited in the ability to encode spatial information by the radio frequency coil array alone which can lead to reduction in signal-to-noise-ratio (Feinberg et al., 2010).

Factors influencing intrinsic DMN connectivity		
Factor	**Issue**	**Publication**
Pulse sequence setting	Temporal resolution	*Smith et al., 2012*
		Feinberg et al., 2010
Scanner hardware	Vendor default choices	*Parker et al., 2014*
	Multisite variability	*Feis et al., 2015*
		Jovicich et al., 2016
2D Multislice acquisition	Slice-order and slice-orientation choices	*Not investigated*
Subject instructions	Eyes-opened / closed / fixated	*Van Dijk et al., 2010*
	Head-motion	*Van Dijk et al., 2010*
Psycho-physiological state	Circadian rhythm	*Blautzik et al., 2013*
	Recreational or medical drugs taken by the subject	*Tanabe et al., 2011*
		Noseworthy et al., 2003
		Rack-Gomer et al., 2009
	Sleep	*Horovitz et al., 2009*
	Human physiology	*Birn et al., 2006*

Table 2.1. Summary table of fMRI acquisition and human physiologic factors influencing the characterization of the DMN.

Although a wide range of sampling rates and a relatively small number of datapoints can be used to measure sufficient BOLD activity for identifying stable spatio-temporal DMN patterns, the standardization of these parameters is still an open issue (Van Dijk et al., 2010; Whitlow et al., 2011). This becomes particularly controversial when harmonizing acquisition sequences across different imaging sites, using different MRI vendors, RF coils and fat suppression methods (Jovicich et al., 2016).

Acquisition methods are defined upon default settings provided by MRI scanner manufacturers which could be different across Siemens, Philips and General Electric

(GE) scanners (Jovicich et al., 2016). These differences also concern 2D slice acquisition methods (Parker et al., 2014) which are usually acquired in an interleaved fashion (Westbrook, 2005) with axial in-plane orientation (O'Connor, 2010). Interleaved slice acquisition is typically chosen to allow for zero or small slice gap while avoiding slice cross-talking effects and related magnetic saturation effects but increases BOLD sensitivity to head motion effects, which are known to bias DMN connectivity (Van Dijk et al., 2012). It was recently proposed that, with some slice gap to avoid slice cross-talking effects, sequential slice acquisition methods might reduce head-motion effects in DMN connectivity (Cheng & Puce, 2014). However, slice-order acquisition choices remain an open issue in the field, especially because these two methods were not directly compared on the same subjects.

As mentioned previously, recent investigations suggest that coactivations in the default-mode system might be influenced also by psycho-physiological states and that individual psychological characteristics should not be underestimated in RS-fMRI designs to characterize consistent DMN connectivity metrics in a sample. In typical designs, subjects are instructed to lay still and relaxing while in the scanner, avoiding the engagement of particular attention towards something in particular and falling asleep (Horovitz et al., 2009). For this reason, it is yet not fully understood whether instructing participants to keep their eyes closed, opened or fixated on a cue could be of help for them to relax without influencing spontaneous brain activity somehow (Van Dijk et al., 2010; Yan et al., 2009). Moreover, considering circadian rhythm effects on the DMN, it would be optimal to scan all individuals at the same time of the day, particularly in longitudinal study designs as in (Meindl et al., 2010). Finally, even though substance abuse is a common exclusion criterion in studies of healthy individuals, the degree of coffee intake (Rack-Gomer et al., 2009), nicotine (Tanabe et al., 2011) or food (Noseworthy et al., 2003) is usually not properly considered and should be at least better supervised.

2.4.2. Image preprocessing issues

After data acquisition, 4D EPI images commonly undergo preprocessing to enhance signal detection and prepare the data for group analysis. RS-fMRI

preprocessing is characterized by a non-standardized workflow of signal processing operations which typically but not necessarily include the following:

- The elimination of some initial TRs to account for magnetic saturation effects.
- Physiological denoising to correct for cardiac and respiratory activities.
- 3D image volume realignment via spatial coregistration.
- Slice-timing correction such that all voxels represent the signal at the same time.
- Spatial smoothing using a Gaussian kernel to increase temporal signal-to-noise ratio (tSNR).
- Temporal smoothing to restrict the analysis to frequencies of interest.
- Regression of motion and additional confounds.
- Mean or mode intensity normalization.
- Spatial normalization to a common space template for group analysis.

The implementation of these operations deserves caution and might be subjected to different ordering to preserve data structure. To improve DMN signal detection, preprocessing workflows include methods to detect and attenuate spurious BOLD fluctuations induced by non-neural biological processes and MRI hardware noise.

2.4.3. Physiological denoising issues

This is not straightforward: the success of these methods is often compromised by the fact that neural and non-neural signals are coupled in time and overlap in space (**Figure 2.4**); therefore their removal will unavoidably be associated with loss in signal of neural origins.

Spontaneous neural BOLD oscillations in the DMN regions occur at very low frequencies (0.01 - 0.1 Hz). In principle, these do not overlap with respiratory (0.1 - 0.5 Hz) and cardiac activity (0.6 - 1.2 Hz) even though DMN signals could also occur at higher frequencies (Cordes et al., 2001). A bandpass filter is typically applied during image preprocessing to remove higher frequencies; however the relatively low sampling rate employed in standard EPI acquisitions (~ 0.5 Hz) does not allow the perfect characterization of these physiological confounds which are aliased in the low frequency range (< 0.1 Hz) (Cordes et al., 2001).

One partial technical solution to this issue consists of recording signals of non-interest such as those rising from human physiology during MRI acqusition to remove them offline by filtering them out of the BOLD timeseries (Glover et al., 2000). MRI-compatible pulse oxymeters and respiration belts can be used to monitor cardiac and respiratory cycles, respectively (Chia et al., 2000; Santelli et al., 2011). However, these parallel measurements cannot often be obtained due to lack of equipment. This issue is especially relevant in multicentric designs where limited availability in parallel measurements across sites is often the case.

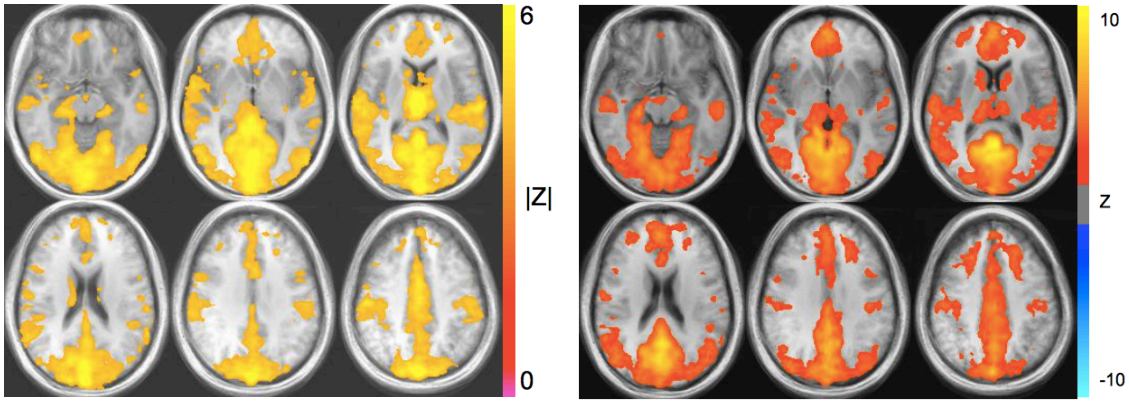

Figure 2.4. Spatial distribution of physiological noise compared with intrinsic DMN connectivity. The figure shows the distribution of physiological artifacts associated with respiration (left panel) and seed-based DMN connectivity (right) averaged across 11 subjects. Physiological activity largely overlaps with DMN coactivation patterns. Adapted from (Birn et al., 2006).

As previously mentioned in the introduction, data-based fMRI methods have been developed to deal with this issue. These are retrospective methods that automatically detect non-neural signal sources from the empirical data. A common approach, the global regression (Fox et al., 2009) removes the whole-brain averaged timecourse from the data using regression methods. This approach assumes that physiological confounds are widespread in the entire brain and that can be consequently catched by averaging the signal across all voxels. A methodological challenge with this approach is that the global average also includes signal from gray matter voxels and therefore unavoidably leads to remotion of signal of interest (Jo et al., 2010). One technical solution in this case includes the adoption of probabilistic segmented anatomical masks of WM and CSF (**Figure 2.5**) to avoid averaging signals from gray

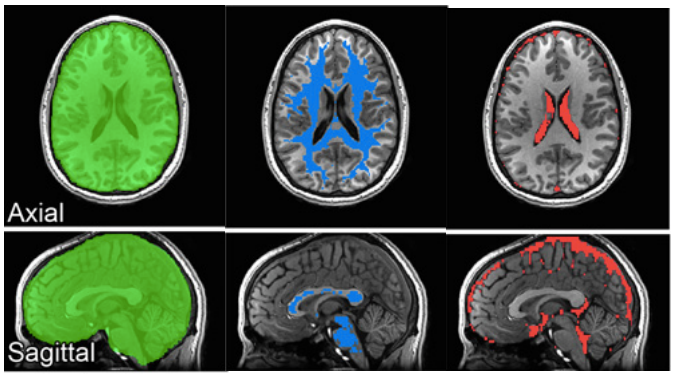

Figure 2.5. Anatomical masks adopted for cardiac and respiratory signal regressions. (green) anatomical masks of the full brain are adopted in global signal regression. Segmented WM (blue) and CSF (red) anatomical masks used to reduce loss in GM neural signals from Jo et al., 2010.

matter (GM) (Weissenbacher, 2009). The global signal regression method remains controversial for inducing anti-correlations in FC-fMRI measures in the entire brain (Fox et al., 2009; Murphy et al., 2009) even though they do not negatively influence DMN connectivity.

Other methods exploit the ICA methodology to detect and separate signal sources associated to physiological noise (**Figure 2.6**). This multivariate approach decomposes the signal into either temporal or spatial statistically addittive independent signal sources by maximizing non-gaussianity or minimizing mutual information between them. Using these methods, human physiology can be observed in the Willis polygon, WM and the ventricles. Despite the relatively low number of timepoints, temporal ICA can be used to isolate temporal modes associated with cardiac (48 - 85 bpm) and respiratory (10 - 24 bpm) frequencies using spatio-temporal priors (Beall & Lowe, 2007).

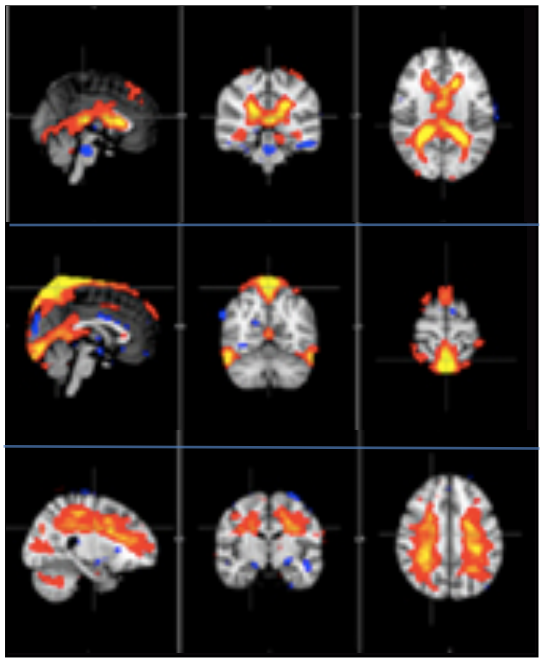

Figure 2.6. Spatial independent components associated with physiological noise. Three known components associated with ventricular activity (top row), vascular activity of veins (middle row) and WM (bottom row). From Tong & Frederick, 2014.

On the other side, spatial ICA methods able to classify and remove various signal artifacts according to spatio-temporal features were also developed (Griffanti et al., 2014; Salimi-Khorshidi et al., 2014). Both methods share ICA limitations and need the user supervision to be conducted. For instance, temporal ICA limitations include the need of a user-defined

frequency range which might be difficult to be standardized across subjects. Spatial ICA classifiers are parametric and might require training the features thus being time-consuming. Furthermore, preprocessing workflows including ICA-based physiological denoising must be defined with attention because slice-timing correction, spatial and temporal smoothing might reduce ICA sensitivity (Beall & Lowe, 2007).

2.4.4. Motion correction issues

If not addressed properly, head motion can severely confound single-subject and group functional DMN connectivity (**Figure 2.7**). Some preprocessing strategies known as *motion correction* or *volume realignment* were developed to deal with motion and are commonly implemented in the majority of resting-state studies. These methods consist of coregistering all train of volumes to a reference volume, by default the volume in the middle of the acquisition train. This procedure shifts all voxels in the exact spatial location defined in the reference volume. Complementary motion correction approaches include the regression of motion metrics obtained from the coregistration process (Satterthwaite et al., 2013) and to censor highly motion corrupted volumes (Power et al., 2012). Similarly to physiological noise correction, ICA-based approaches were developed to automatically detect and correct for head-motion (Pruim et al., 2015a; Pruim et al., 2015b; Schopf et al., 2010; Schopf et al., 2011; Wang et al., 2012).

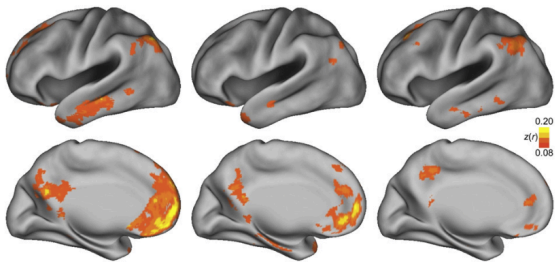

Figure 2.7. Surface maps revealing differences in head-motion in the main DMN regions. Differential group seed-based intrinsic FC-fMRI maps illustrate how head motion might confound DMN connectivity. Each map represents the FC-fMRI difference with lesser motion as compared to a second group with greater motion. The leftmost image shows the contrast between the two most extreme groups. The rightmost image shows the contrast between groups that have a subtle difference in mean motion estimates. Even tight differences in head motion yield difference maps that could easily be mistaken for neuronal effects. Adapted from (Van Dijk, et al., 2012).

For several reasons, the implementation of motion correction is always recommended even if motion is overall low. First, it has been show that even

submillimetrical motion can confound DMN connectivity estimates (Power et al., 2012). However, if motion estimated is low, ICA methods can quite efficaciously deal with it (**Figure 2.8**) and therefore motion correction could be avoided to preserve data structure (Meindl et al., 2010). Second, the order of slice-timing and head-motion corrections cannot be predefined but depends on slice-order acquisition methods and the amount of motion detected in the data. Some methods which can perform both corrections simultaneously have been proposed (Roche, 2011) while other argued that applying motion correction prior slice-timing correction would be an optimal solution (Jones et al., 2008) especially if some preprocessing software and algorithms require slice-timing information to be implemented (Beall & Lowe, 2007). In general, relative freedom of choice is admitted for interleaved slice acquisition protocols. Instead, different optimal solutions are proposed in case slices are acquired sequentially as a function of head-motion severity (Huettel et al., 2008; Poldrack et al., 2011).

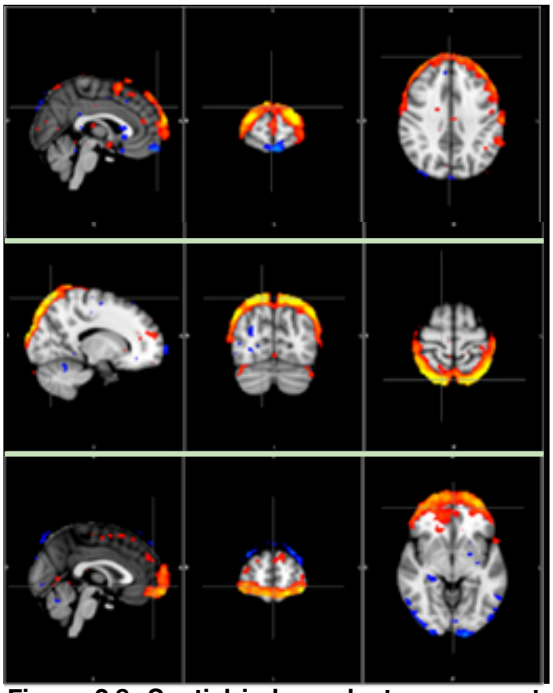

Figure 2.8. Spatial independent components associated with head-motion. The figure shows three typical motion-related spatial components identified using ICA on preprocessed data. From Tong & Frederick, 2014.

The relevance of slice-order acquisition methods for head-motion correction was already described previously among the fMRI data acqusition issues. Regarding this issue, volume realignment is only a part of motion correction. Volume-based motion correction methods treat brain volumes as a rigid-body and perform inter-volume coregistration. While the rigid-body assumption is technically wrong (Poncelet et al., 1992), these methods do not correct for motion occurring within-volume associated with timing differences in the acquisition of consecutive 2D slices ($0 < t < TR$). More recently, advanced retrospective motion correction solutions have been proposed with the introduction of novel slice-wise correction methods (Beall & Lowe, 2014). In particular it has also been showed that standard volume-based methods might be quite

insensitive to motion in BOLD-weighted MRI data (Beall & Lowe, 2014). In any case, it remains unclear how within-volume motion occurring during different slice-timing acquisitions might influence the intrinsic DMN connectivity (Kim et al., 2008; Sladky et al., 2011). It should be noted, however, that no methods can account for motion in its entirety so far, with slice-based motion correction (SLOMOCO) slightly outperforming the others.

2.4.5. DMN extraction methods

Following preprocessing steps, several methods have been developed to characterize intrinsic DMN connectivity. These can be grouped into two categories here defined as model-based or model-free methods. Model-based methods are hypothesis-driven approaches which exploit predefined anatomical region(s)-of-interest (ROIs) to investigate intrinsic FC-fMRI in the entire brain. They permit to gain an immediate inspection and high interpretability about the intrinsic human brain functional architecture. In contrast, model-free methods are exploratory and do not require the predefinition of anatomical priors to characterize intrinsic brain networks (Beckmann et al., 2005; Calhoun et al., 2001a). Both methods present advantages and disadvantages and their choice strictly depends on the scientific question of the experimenters.

The simplest and most common model-based method is the seed-based approach. This method characterizes FC-fMRI by means of univariate temporal correlations between the averaged signal within a predefined ROI, technically referred to as the *seed*, and the signal in all other voxels/ROIs throughout the brain (Fransson, 2005). Another model-based method exploits graph-theory analysis to investigate topological properties of intrinsic brain networks. This method is a mathematical representation of brain networks as graphs, characterizing FC-fMRI as *edges* or *archs* between a set of predefined ROIs, technically referred to as the *nodes* of the graph. Topological properties of intrinsic brain networks include *clustering coefficien*t i.e. the degree of local interconnectivity between nodes; the *path length*, i.e. the efficiency of the neural transmission among distant nodes; *hub analysis* to identify those ROIs that plays a fundamental role within the network, et cetera (Bullmore & Sporns, 2009; Hosseini & Kesler, 2013). Both methods are used to identify the DMN by placing the *seed* in the posterior cingulate cortex and the precuneus, considered the main hub of the

DMN (Fransson & Marrelec, 2008; Leech et al., 2012; Leech et al., 2011). The definition of the shape and size of ROIs poses serious challenges to characterize intrinsic DMN connectivity (**Figure 2.9**), considering eventual parcellation scheme (Shen et al., 2013) and, edge thresholding in graph-theory approaches (Van Wijk et al., 2010).

Model-free methods include multivariate approaches such as ICA (Bell & Sejnowski, 1995), an exploratory method that decomposes the BOLD signal into its addittive, statistically independent, spatio-temporal sources. Therefore, ICA can automatically identify independent spatio-temporal sources of resting-state brain networks, including the DMN, in addition with the aforementioned artifacts associated with head-motion and physiological noise (Kiviniemi et al., 2003; McKeown, 1997). ICA cannot be perfomed on space and time simultaneously. For this reason, although it is possible to identify the DMN by conducting ICA in either dimensions (Boubela et al., 2013), the spatial dimension is preferred because of the larger number of voxels compared with time points (Beall & Lowe, 2007).

Moreover, different ICA approaches do exist. The first approach consists of conducting an ICA on individual EPI data. This method is usually named single-subject ICA. An alternative approach consists of conducting an ICA on the entire group of subjects, concatenating all the EPI volumes in time (Calhoun & Adali, 2012). Then, methods to characterize the activation patterns in each individual include back-reconstruction (Calhoun et al., 2001b) or dual-regression (Beckmann et al., 2009). Very recently, it has been shown that group ICA approaches improve separability of neural from artifactual signals improving the ultimate ICA sensitivity to FC-fMRI signals (Du et al., 2016).

Despite its advantages, ICA also comes with some limitations. First, ICA is non-deterministic which implies that the ultimate decomposition of the BOLD signal sources always allows some degree of run-to-run variability. Second, ICA is also unsupervised which implies that no predefined number of components exists and that this could be therefore extablished using semi-quantitative methods or empirically by (re-)conducting ICA at different dimensionality orders on the same data. Suboptimal dimensionality order might either cause signal overlapping (erring on the side of caution) or splitting across several components (erring on the side of excess) (**Figure 2.10**).

In the former case, DMN signals will coexist with irrelevant signals in the same component. In the latter, DMN signals will be split across several components introducing spurious interregional variability in connectivity patterns within one single component. Similarly, if spatio-temporal signal patterns largely overlap, ICA might fail to separate independent signals sources. This would introduce some degree of

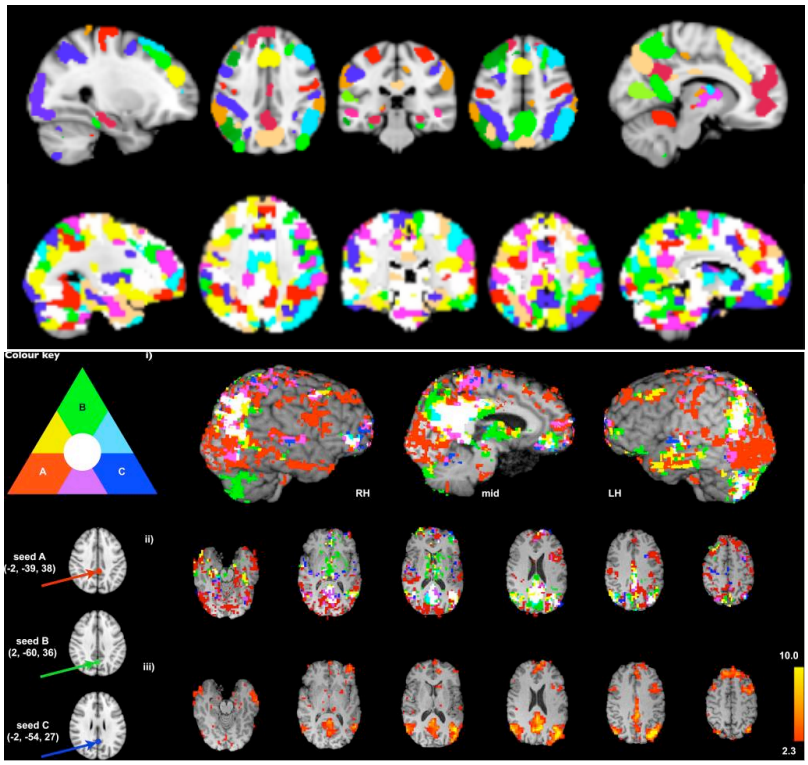

Figure 2.9. ROI definition issues and Intrinsic DMN connectivity as a function of seed location. The top figure shows two atlases made of 90 (Shirer et al., 2012) and 499 functional ROIs (Richiardi et al., 2015) respectively, while the bottom figure shows the DMN characterized from three identical seeds located in different positions. Seed voxel locations from (Fox et al., 2005) in red; (Singh et al., 2008) in green; (Greicius M. D., 2003), in dark blue; Model-based methods are limited in the definition and location of seeds or nodes, atlases or parcellation schemes which lessens repeatability and reproducibility across studies. © 2010 Cole, Smith and Beckmann.

interdependence among components leading to the imperfect characterization of the DMN or other resting-state networks, which will be likely to be contaminated by non-neural sources (Beall & Lowe, 2007). Overall, these issues might lead to reduce DMN consistency across individuals and reduce the reproducibility of intrinsic fMRI studies (Schopf et al., 2010).

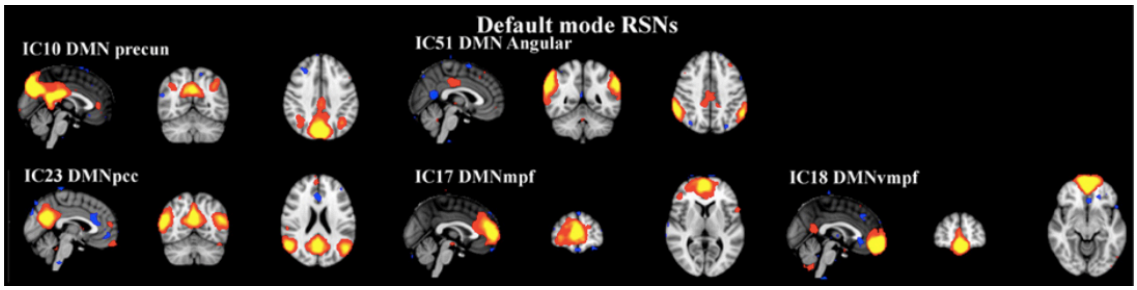

Figure 2.10. The DMN splitting phenomenon following ICA. The figure shows an example of strong DMN splitting onto 5 independent components following group ICA. From (Rytty et al., 2013).

2.5. Intrinsic DMN connectivity: a potential biomarker?

Many definitions of biomarkers have been proposed over the last decades. Fortunately they converge quite well with the general definition given by the National Institutes of Health (NIH). According to this definition, a biological marker or biomarker is a metric that can denote both normal biological and pathogenic processes or pharmacological responses to therapeutic drugs (Biomarkers Definitions Working, 2001). Neuroimaging biomarkers are detailed biological features derived from radiological images or recorded brain signals that can be used in isolation and jointly with known biomarkers to assess or predict the presence of disease and evaluate treatment response (Richter, 2006).

Diagnostic biomarkers can be clinical, i.e. able to assess the presence of disease, and preclinical, i.e. able to assess signs or risk factors for health and function deterioration (Shoji, 2012). Preclinical neuroimaging biomarkers are particularly important in case of neurodegenerative diseases or cancer that present a symptom-free phase. In these cases, these biomarkers might timely identify high-risk individuals so that they receive either preventive treatments or further follow-up evaluations (Frank & Hargreaves, 2003; Sapsford et al., 2010). This class of biomarkers include hippocampal atrophy measured in morphometric MRI studies of Alzheimer Disease (AD) (Fraser et al., 2015). This MRI-based biomarker can be used in conjunction with CSF analytes that measure abnormal protein aggregates (i.e. amyloid beta, phosphorilated tau protein) or with indicators of neural loss to distinguish between disease-modifying and symptomatic treatment effects (Hampel et al., 2011; Saumier et al., 2009).

Other relevant longitudinal biomarkers of AD are obtained from molecular neuroimaging using PET. The Pittsburgh compound-B (PiB-PET) tracer and Fludeoxyglucose PET (FDG-PET) represent validated biomarkers of amyloid-beta plaques accumulation and abnormal glucose metabolism, respectively (Frisoni et al., 2013; Habeck et al., 2012). Nevertheless, healthcare providers resist compensating for the expensive costs of these diagnostic biomarkers for insufficient clinical evidence about diagnosis or treatment of illness and improvement of cognitive functioning in people affected by dementia (Frisoni & Visser, 2015).

While technique noninvasiveness, ease of acquisition and unexpensive costs are standard criteria to qualify preclinical biomarkers (Katz, 2004a, 2004b), further evidence from comparative and longitudinal studies involving subjects from appropriate populations is demanded (Frisoni & Visser, 2015). In contrast, these neuroimaging biomarkers reflect neurodegeneration and metabolic dysfunction only in symptomatic stages of AD but not in its preclinical asymptomatic stages where structural MRI and FDG-PET efficacy is lessened while PiB-PET becomes more challenging considering that amyloid-beta deposition is not AD specific and may be also found in normal aging (Jack et al., 2012; Toledo et al., 2015).

2.5.1. Intrinsic DMN fMRI-based biomarker

Intrinsic DMN connectivity as defined by fMRI offers attractive possibilities for evaluation as a biomarker for some diseases. Evidence from cross-sectional studies suggets that the intrinsic DMN connectivity is sensitive to neuropsychiatric illness and neurodegenerative diseases showing aberrant loss in connectivity strength with regional dysconnectivity or hyperconnectivity. Since BOLD-weighted FC-fMRI is not a direct measure of neural activity, these alterations likely indicate metabolic dysfunction that would cause an incomplete recruitment of DMN regions or induce additional brain effort to safeguard network integrity and support cognition with higher computation costs (Palop et al., 2006).

DMN-based biomarkers can be easily developed considering that the network is consistently observable in healthy young adults (Damoiseaux et al., 2006) and that default-mode connectivity predicts cognitive performance (Sala-Llonch et al., 2012). The ability to characterize the DMN in young healthy adults has allowed the study of its

neurophysiological changes (Chou et al., 2012; Patriat et al., 2013; Shehzad et al., 2009) and to characterize abnormal DMN connectivity changes in pathological conditions (Somandepalli et al., 2015; Song et al., 2012). This is of primary importance for studying typical and atypical development in children (Uddin et al., 2010) where aberrant connectivity patterns undergoing atypical brain maturation trajectories were identified (Uddin et al., 2010). These are found in the DMN that show a retarded formation of distant connectivity in autism (ASD) until late adolescence (Washington et al., 2014).

There is evidence supporting DMN-based biomarkers of major depression which causes changes in anterior-posterior DMN connectivity with increased connectivity in anterior DMN regions (Craddock et al., 2009; Greicius et al., 2007; Mulders et al., 2015; Sheline et al., 2009). Increases in intrinsic DMN connectivity are also typically found in case of structural white matter degeneration in multiple sclerosis (MS) probably because of compensatory mechanisms that preserve deficitary cognition (Hawellek et al., 2011).

Diminished FC-fMRI can instead be used as a DMN-based biomarker of normal aging and more relevantly for mild cognitive impairment (MCI) and mild AD (Balthazar et al., 2014; Greicius et al., 2004). DMN connectivity shrinkage in normal elderly is a natural process associated with aging decline and gray matter atrophy which does not necessarily indicate the presence of disease. However, evidence suggests that DMN connectivity loss is related to increased amyloid-beta (Hedden et al., 2009; Sheline et al., 2010) and symptoms' severity of dementia (Hafkemeijer et al., 2012). These findings consequently promoted the ability to track and monitor clinical deterioration in MCI and AD patients (Binnewijzend et al., 2012; Damoiseaux, 2012) and discriminate between AD and other forms of dementia such as fronto-temporal dementia (FTD) (Zhou et al., 2010) or dementia with Lewy bodies (Galvin et al., 2011). Importantly, intrinsic DMN connectivity might distinguish between MCI who would undergo cognitive decline, hence conversion to AD, from those who would remain more stable over time (Prestia et al., 2013; Vos et al., 2013).

Longitudinal task-based BOLD fMRI studies provided evidence in support of DMN-based predictive biomarkers of dementia (Petrella et al., 2011) but longitudinal RS-fMRI studies are still exiguous in this context (Drago et al., 2011) because of little

knowledge in the reliability of the default-mode system (Zuo & Xing, 2014). This not just a case. Physiological and head-motion related artifacts that are not sufficiently removed from the BOLD signal using optimized acquisition protocols associated with tailored preprocessing strategies induce changes in intrinsic DMN connectivity patterns that can be misunderstood as disease or treatment related effects. TRT reliability investigations are important in this context because it has been shown that non-neural BOLD components will reduce DMN reliability even when the tSNR would indicate low noise in the data (Mueller et al., 2015). Therefore, TRT reliability studies would be useful to investigate the stability of DMN connectivity measures over time and optimize longitdunal RS-fMRI protocols regarding acquisition and analysis methods implemented.

Different methods exist to quantitatively assess the TRT reliability of the DMN. Some methods are complementary but not necessarily interchangable, depending on sample size or mathematical constraints of some FC-fMRI metrics. These are better outlined in the following paragraph.

2.5.2. Test-retest reliability methods

In general reliability refers to the degree of stability exhibited when a measurement is repeated under identical conditions. In research, reliability analyses are conducted to evaluate the extent to which unsupervised sources of noise introduce variability in the measurements obtained from a given scientific apparatus. As such, reliability analyses are conducted under the assumptions that measurements are valid and are not changing across repeated conditions. In the context of resting-state brain MRI experiments with human subjects, there are many factors that might introduce error while measuring brain function. Some depend on the MRI hardware (i.e. thermal noise) or task setup (i.e. clarity of instructions, personality of the examiner) and some on individual characteristics of the participant (i.e. health, fatigue, coughing, motivation, momentary distractions, comprehension of task).

Reliability becomes of particular interest when evaluating the consistency of measurements obtained from the same subject over multiple trials over time. This TRT reliability method quantifies the degree of intra-individual variability and is pratically used in medical contexts to monitor patients and differentiate the relative contribution

of disease or treatment from confounding factors in the measurement change. In the MRI context, TRT reliability is appropriate if scanning repetitions occured over a short time and were conducted by the same experimenter at the same MRI center using the same MRI scanner, image acquisition protocol and data analysis software.

Reliability can be measured using different mathematical and statistical approaches. A general approach includes standard pairwise correlation methods using parametric Pearson's or non-parametric Spearman's correlations between all measurements across sessions. Correlations yielded will fall in the range [-1, 1] with values higher than 0.5 indicating a moderate to high degree of test-retest reliability, values in between 0 and 0.5 indicating a low degree of reliability and negative values indicating unreliable measurements. Although intuitive, this approach was not adopted in the reliability studies reported in this work because more sophisticated methods have been introduced recently.

The TRT reliability method adopted in these longitudinal studies is a measure of absolute percent change in the measurements across repetitions for each single subject. In the present work the absolute percentage change was mathematically defined as the fraction between the absolute difference and the mean of measurements across repetitions expressed in percentage. This TRT reliability measures the error across repetitions ranging from 0 to ∞. Since the measurement error and reliability are indirectly proportional, the more the error approaches 0 the higher the reliability and vice versa. Of note, this reliability measurement is unitless and needs thresholding definitions to allow interpretations and decisions. In the present work, 10% error across repetitions was conventionally defined as the limit between good and poor reliability. Also, this method is not adequate for examining the reliability of seed-based pearson's correlation coefficients unless thresholding rho $> \pm 0.1$, to avoid zero correlations that bias reliability estimates toward ∞. Moreover, for similar formulation reasons, higher correlations will be less reliable than lower correlations. Since Fisher's z-transformation typically increases pearson's correlation coefficients, this will lead to reliability loss. In contrast, this method can be efficaciously implemented on ICA z-scores values that are usually thresholded at $z > \pm 2.3$, $p < .001$.

Other reliability methods include inter-rater approaches which are well suited in multicentric fMRI studies to evaluate the degree of absolute agreement across different

MRI centers, MRI machines and analysis protocols on different individuals. However, inter-rater methods can also be adopted in single site studies as a measure of concordance of measurements across repetitions in all individuals. To this extent, inter-rater methods can be adopted along with the absolute percent change to circumvent its limitations.

Inter-rater reliability is calculated via ICC-analysis (Shrout & Fleiss, 1979). This coefficient is mathematically inscribed in an analysis-of-variance (ANOVA) framework and quantifies the "consensus" as the variability measured within all subjects and between repetitions (concordance) and/or across multiple MRI sites (agreement). Simplifying, ICC can be therefore calculated as the fraction of between-subject variance and the total variance detected in the sample (Bennet & Miller, 2010). This formula allows a range for ICC values between 0 and 1 with moderate-to-high reliability in the range [0.5, 1] and poor reliability otherwise. Furthermore, there are different types and models of ICC depending on whether reliability measures are calculated from single or multiple averaged measures (type 1 and type 2, respectively) and, depending on the study design, if the same subjects were considered across conditions or if they were randomly selected for each condition. The main difference between the ICC and the absolute percent change regards the fact that the former computes reliability in a group-wise fashion while the second can be used to monitor variance in each single subject or voxel. Since the ICC is mathematically derived from the ANOVA it could be similarly biased by outliers or underpowered in small samples.

In the main works outlined in this thesis the absolute percent change and the ICC methods were both adopted to investigate the stability of DMN connectivity measurements. This choice was in part related to the study design of these longitudinal studies and to their popularity in the current investigations in the field.

Furthermore, other reliability measures adopted in these works include the Jaccard and Dice indexes. These are statistics that compare the similarity of two or more samples (Tan et al., 2005) and have been adopted in reliability studies of brain imaging to evaluate voxel-wise spatial reproducibility of brain network patterns across sessions (Rombouts et al., 1997; Rombouts et al., 1998). Jaccard and Dice indexes are very similar statistics with Jaccard = Dice / (2 - Dice). The Jaccard index is computed from the ratio between the intersection and the union of the two or more activation patterns whereas the Dice index is

computed as the ratio between the intersection and the mean spatial pattern between sessions. Both statistics return values in-between 0 and 1 indicating poor and high reliability, respectively.

2.6. DMN reliability: a systematic review

TRT reliability studies are fundamental to optimize longitudinal RS-fMRI protocols and evaluate the clinical value of intrinsic DMN connectivity. To date, despite efforts in the scientific community, our knowledge about DMN reliability remains limited and need further investigations. Overall, the existing studies suggests that functional DMN connectivity metrics are overall substantially stable in the long term, suggesting that this network can be a promising candidate to validate biomarkers in longitudinal RS-fMRI studies. However, TRT reliability metrics can be influenced by MRI acquisition and preprocessing choices. In the next paragraphs, some of known effects are illustrated along with open issues to be addressed.

2.6.1. Longitudinal data acquisition issues

Despite large variability in MRI acquisition protocols, TRT reliability of the DMN is generally substantial to support the development of novel biomerkers. High long-term DMN reliability is indeed found in 1.5T studies as well as at higher field strength (Chou et al., 2012). Notwithstanding the success to detect the DMN at ultra high-field (Hale et al., 2010), no studies have so far either evaluated DMN reliability at 7T or compared it across different field strengths.

Most reliability investigations have been conducted at single MRI centers (Chou et al., 2012; Guo et al., 2012; Patriat et al., 2013; Shehzad et al., 2009; Somandepalli et al., 2015). These studies may vary in EPI acquisition schemes which can be determinant for ultimate reliability metrics, such as number of acquired volumes (Braun et al., 2012), limiting reproducibility and clinical relevance of findings. Longitudinal datasets are unique and often challenging to obtain, so their public availability may help generalizing the results from different analyses methods. Within, the 100 functional connectome project (http://fcon_1000.projects.nitrc.org/), the Consortium for Reliability and Reproducibility (Zuo & Xing, 2014) is a recent effort that created a public

repository of various longitudinal RS-fMRI datasets independently collected worldwide on 3T MRI machines (Siemens, Philips, GE) using a variety of MR acquisition protocols. Another important neuroimaging data-sharing initiative is the International Consortium for Brain Mapping (ICBM) from which interesting DMN reliability information have been gained (Song et al., 2012). Other efforts to build 3T multisite datasets come from The PharmaCog project (Jovicich et al., 2016) that collected data using a harmonized protocol from a European consortium of 13 MRI centers under a large umbrella of hardware characteristics that span from scanner manufacturer (Siemens, Philips, GE) to head-coils (8-12 channels) and fat-suppression methods (constrained to vendor's choice).

As mentioned in previous sections, subjects instructions might differently influence FC-fMRI estimates. This issue of eyes-condition is particular relevant for DMN reliability studies since it cannot be optimized apriori. While being disregarded in some studies (Guo et al., 2012; Song et al., 2012) and in disagreement across studies (Chou et al., 2012; Shehzad et al., 2009), some have directly addressed this issue (Patriat et al., 2013). Their conclusion was that eyes-fixated would provide more reliable connectivity estimates. However, given the complexity of the field further investigations are needed to corroborate this solution.

Despite the strong dynamical nature of the default-mode function, all FC-fMRI measures provides high reliability within DMN regions. However, biological changes in DMN function remain an open issue, especially in clinically oriented longitudinal studies that implement long-term repetitions to mimic clinical trials (Chou et al., 2012; Jovicich et al., 2016). However, a conservative approach would consider also short-term repetitions in clinical longitudinal studies such as in (Somandepalli et al., 2015) or in multisession (i.e. more than 2 scanning repetitions) studies (Shehzad et al., 2009) to improve the statistical power of reliability analyses.

An important source of variability across longitudinal acquisition protocols is related to the number of 2D slices adopted to image full 3D volumes. Most importantly, no studies either reported or compared slice-order acquisition sequences in longitudinal studies, which does not help clarifying whether interactions between head-motion and slice-timing acquisition differences would influence the TRT reliability of the DMN.

Finally, considering that it has been recently shown that DMN connectivity is sensitive to circadian rhythm in healthy young adults, with gradually reduced synchronization from morning to afternoon (Blautzik et al., 2013), and that rhythmicity dysregulation undergoing in clinical populations would affect DMN connectivity (Blautzik et al., 2014), TRT reliability could be higher following standardization of acquisition time, particularly in the morning, when DMN connectivity is the strongest (Meindl et al., 2010).

2.6.2. Longitudinal data analysis optimization

Longitudinal RS-fMRI studies do not show significant changes from standard single-session designs. Indeed standard correction procedures and FC-fMRI methods allow the characterization of stable connectivity metrics over time. In 1.5T fMRI protocols, minimal preprocessing consisting of motion correction and bandpass filtering was able to return high multisession seed-based DMN reliability, with 45% of DMN connections showing outstanding ICC scores (ICC $\geq$ 0.8) (Chou et al., 2012). In 3T studies, which constitutes the majority in the literature, many attempts were made to reduce spurious fluctuations rising from non-neural signals that would lessen TRT reliability of FC-fMRI metrics.

Besides common physiological denoising strategies including WM/CSF and global signal regression, (Birn et al., 2014) compared longitudinal effects of RETROICOR (Glover et al., 2000), progressively adding respiration-related metrics (Bianciardi et al., 2009; Birn et al., 2006), relative derivatives (Bianciardi et al., 2009), cardiac rate correction metrics (Chang et al., 2009) with NPC on model-based FC-fMRI in healthy young adults. In contrast to their expectations, they found that physiological noise corrections determined significant decreases in TRT reliability in the whole brain. In agreement with (Braun et al., 2012; Guo et al., 2012; Song et al., 2012), the greatest decreases were found for those denoising approaches using nuisance regressors derived from the data itself, specifically global signal regression (Birn et al., 2014). However, when only brain networks ROIs were interrogated, the results were similar, except that the WM/CSF, and global signal regression did not result in a significant difference in ICC values (Birn et al., 2014). This counterintuitive effect might be explained considering that physiological noise can be characteristic and repeatable within subjects,

and reduction of this noise would therefore reduce TRT reliability of individual FC-fMRI metrics (Birn et al., 2014).

These observations are indeed surprising and do not easily lead to the conclusion that physiological noise correction should not be performed in longitudinal RS-fMRI studies. Rather, physiological noise correction would constitute a tradeoff between reliability and specificity of connectivity results. Even so, further investigations involving healthy old participants who show higher between-subject variability in respiratory and cardiac rates might confirm these findings. Furthermore, (Birn et al., 2014) did neither specifically interrogate DMN regions nor compare different FC-fMRI methods.

Particular attention was also given to the temporal filter frequency range in the context of model-based whole-brain intrinsic FC-fMRI. Typical RS-fMRI preprocessing use bandpass filters in the range 0.01 - 0.1 Hz. When comparing a narrower frequency range (0.04-0.08 Hz) with a broader one (0.008-0.15 Hz), a larger range returned higher TRT reliability in the whole brain (Braun et al., 2012). These choices have not been investigated in the context of ICA-based methods but are particularly interesting in the context of fast sampling rate acquisitions that allow the investigations of higher frequencies ranges, spare of physiological artifacts, eventually leading to the discovery of neural network hubs (Liao, 2013).

Certainly, global signal regression received particular attention in DMN reliability studies. As described previously, this method introduces anticorrelations in the entire brain that would overall reduce model-based reliability in healthy young and old adults up to 50% (Braun et al., 2012; Guo et al., 2012; Song et al., 2012). However, the severity of reliability loss seems more pronounced in seed-based and ROI-based connectivity than graph-based approaches (Braun et al., 2012). Auspiciously, global signal regression is less than an issue in longitudinal DMN studies since anti-correlations are not introduced within the DMN regions (Guo et al., 2012).

2.6.3. Comparison of functional connectivity methods

The majority of TRT reliability studies of the DMN was conducted using model-based connectivity methods, particularly seed-based or ROI-based approaches. These manifest overall fair reliability estimates in healthy young individuals, ICC (0.47 ±

0.16; range 0.3 - 0.69) without global signal regression (Braun et al., 2012). However, the magnitude of the correlations affects the ICC analysis with higher between-subject than within-subject variance (Birn et al., 2014; Chou et al., 2012; Shehzad et al., 2009).

Besides, the main issue with model-based connectivity approaches relates to the selection of *seeds* and *nodes*. Considering the high variability of FC-fMRI maps as a function of seed location and characteristics, PCC-located seeds (MNI coordinates: –6 – 58 28) manifest high DMN reliability only at short-term (Shehzad et al., 2009). One solution to optimize ROI definition can be the adoption of ROI atlases. The automated anatomical labeling (AAL) atlas would improve DMN reliability in the long-term (Wang et al., 2011). No differences in DMN reliability were observed in short-term reliability across commonly adopted anatomical parcellation templates (Desikan et al., 2006; Dosenbach et al., 2010; Toro et al., 2008).

Optimizing ROI choices is important to evaluate seed-based DMN reliability differences across healthy subjects and patients. When comparing typical developing with attention-deficit/hypercativity disorder (ADHD) children, seed-based connectivity ICC is the highest in regions exhibiting significant correlations with PCC (ICC > 0.4), but across the whole brain ranged from fair to moderate (ICC, 0.2 – 0.6) (Somandepalli et al., 2015). These non-zero ICCs found outside the DMN indicate the existence of reliable within-subject noise that likely confounds group comparisons (Patriat et al., 2013). To further reduce these biases, ICA-informed ROIs could also be used to optimize individual ROIs definition in the DMN (Franco et al., 2013). However, this would need for a direct comparison between the TRT reliabilities of model-based and model-free connectivity methods within the default-mode system.

Intrinsic issues related to model-free methods, such as DMN-splitting, lessen ICA-based DMN reliability (Jeong et al., 2012; Meindl et al., 2010) but could be circumvented by determining low-order dimensionality (Abou-Elseoud et al., 2010). Despite being not appropriate to investigate the functional connectome to large extents, the decomposition of groupwise data into 10 components proved successful at characterizing highly reliable multisite DMN connectivity patterns in healthy aging individuals (Jovicich et al., 2016). Group-ICA methods produce more reliable DMN connectivity estimates compared with single-subject ICA methods (Guo et al., 2012;

Zuo et al., 2010) and their application is favored in longitudinal study designs given that multiple acquisitions increase group-ICA sensitivity (Franco et al., 2013).

DMN reliability studies are relevant in the context of healthy aging considering age-related reliability differences, with healthy elders showing poorer reliability when compared to young adults (Song et al., 2012). This is an interesting finding that can lead to the development of DMN-based longitudinal biomarkers or multivariate classification algorithms used to distinguish age-related or abnormal FC-fMRI changes (Meier et al., 2012).

To strengthen this research stream, (Jovicich et al., 2016) recently showed that in the context of 3T multicentric studies, despite variability in the tSNR across multiple sites due to hardware and pulse-sequence differences, both seed-based and ICA-based DMN connectivity methods showed consistent TRT reliability across 13 different MRI sites with the latter methodology providing overall higher DMN reliability.

This result agrees with previous single-site DMN reliability studies of healthy aging (Guo et al., 2012) and could depend on technical differences or different degrees of sensitivity to motion artifacts between the two FC-fMRI approaches (Guo et al., 2012). In the latter study, seed-based and ICA-based approaches were compared as a function of across-session head-motion severity (**Figure 2.11**).

In this study, motion metrics calculated during 3D volume realignment were used to separate the entire sample (N=24) into two subgroups: ‘mvmt<3’ (N = 20) and ‘mvmt<2’ (N = 15) that included subjects with less than 3 mm and 2 mm of movement fluctuation in either translation or rotation across sessions, respectively. This study indicates that both methods agree in good large-scale and poor voxel-wise DMN reliability, with a slight advantage for model-free methods at the voxel-wise level. In contrast to model-based approaches, ICA-based DMN reliability is more robust against head-motion in healthy aging (Guo et al., 2012). However, as previously described, volume-based motion correction methods account for motion only in minimal part (Beall & Lowe, 2014) and no longitudinal study has so far compared different retrospective approaches to minimize motion correction effects on model-based and model-free DMN reliability.

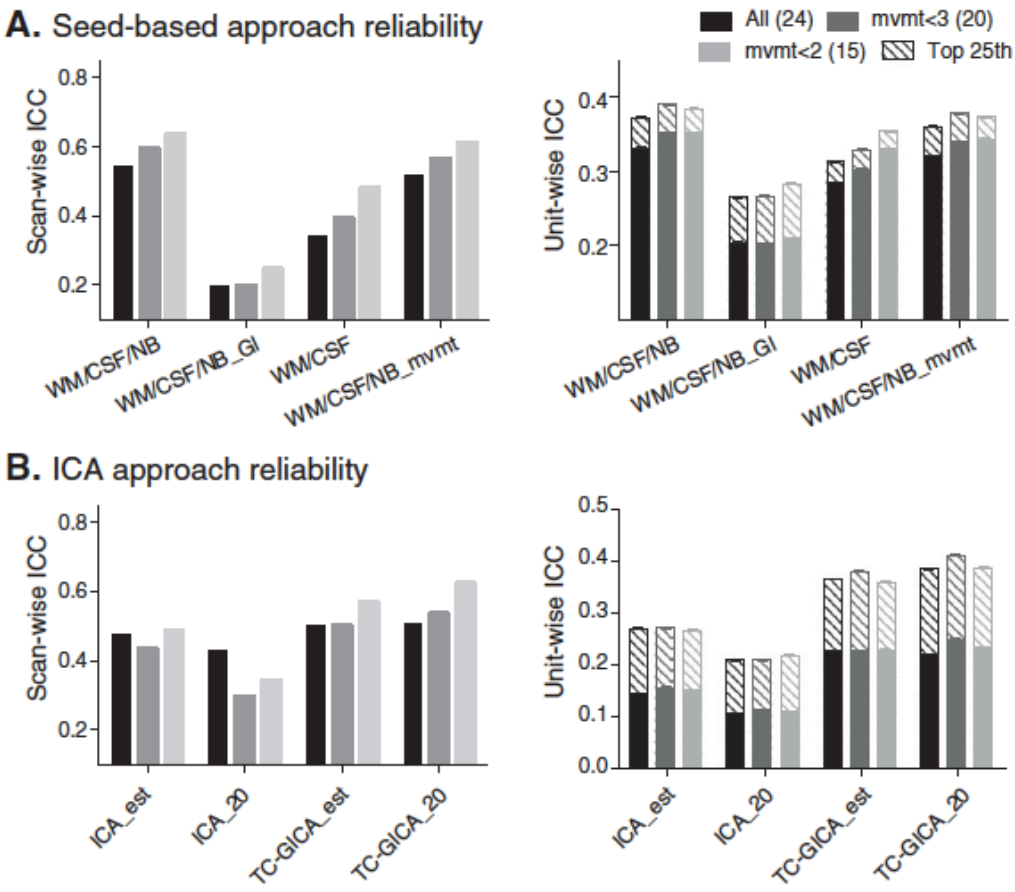

Figure 2.11. Comparison of seed-based and ICA reliability analysis using ICC. A) shows seed-based ICC analysis and B) shows ICA-based ICC analysis at the ROI (left) and voxel-wise level (right), respectively. Bargroups include all individuals (black), high-motion individuals (dark gray), low-motion individuals (light gray). The image shows that global-regression induce loss of seed-based reliability and that remotion of motion parameters or non-brain signals including WM/CSF can improve reliability. Group-ICA returns higher reliability than single subject ICA approaches. Higher reliability is found in large ROIs compared with single voxels. However, ROIs are more susceptible to motion. WM/CSF/NB: regression of WM/CSF and non-brain tissue; WM/CSF/NB_GI: includes global regression; WM/CSF: use only WM/CSF regressors; WM/CSF/NB_mvmt: includes motion parameters. ICA_est: single-subject ICA with automatic dimensionality estimation. ICA_20: single-subject ICA with 20 components; TC-GICA_est: group ICA with automatic dimensionality estimation; TC-GICA_20: group ICA with 20 components. From Guo et al., 2012.

2.7. Challenges addressed in this thesis

This thesis uses TRT reliability metrics of the intrinsic DMN connectivity as a probe in two separate studies that investigate different strategies for physiological noise (Chapter 3) and head motion correction (Chapter 4) methods. Both studies used group-

ICA and dual-regression methods to define DMN and to capitalize on ICA's power to eliminate unwanted sources of signal noise. The general assumption is that acquisition and/or analyses methods that improve TRT reliability (i.e., reduce test-retest variability), should be preferred in longitudinal studies evaluating the value of DMN as a biomarker.

3

Experiment 1: Influence of physiological denoising on multisite DMN reliability in healthy aging

3.1. Introduction

The first challenge addressed in the present thesis relates to physiological aliasing and correction methods in longitudinal RS-fMRI studies of the DMN. The purpose of this study was to evaluate whether different rPNC approaches, intended as methods that minimize the cardiac and respiratory components in the resting state BOLD fMRI signal, would differently affect the within-site TRT reliability of DMN connectivity across two MRI sessions and the consistency of the reliability in healthy aging subjects from the PharmaCog consortium. In this 3T harmonized multisite fMRI study, three ICA-based physiological denoising algorithms were evaluated, namely PESTICA and FSL-FIX along with the WM/CSF regression method. These were compared to a control condition where no retrospective physiological denoising was implemented (NPC). The raw data of this study is public so it can be used to evaluate additional physiological or head motion correction methods.

The rationale of the current study is to extend previous findings (Birn et al., 2014) to 3T multisite studies, healthy aging individuals and model-free intrinsic FC-fMRI methods. To the best of our knowledge it remains unclear how rPNC may affect TRT reliability in the context of a multicenter MRI study. In the context of multisite experiments, rPNC methods are of special interest because short TR acquisitions or direct physiological measurements are more challenging to implement uniformly across clinical sites that may not have special equipment or access to special MRI sequences.

In particular, reliability (or lack thereof) may be exacerbated in heterogeneous populations such as healthy elderly subjects. In the context of the current study, this age group is relevant because the patterns of physiological noise variability might not be the same in young and elderly people (Nicolini et al., 2012; Schulz et al., 2013; Taylor, 2010). The contributions of cardiac and respiratory activity on BOLD signal fluctuations as a function of age are unknown, so a conservative approach is to consider that physiological effects in young and elderly people may be different. Therefore, evaluating how physiological correction methods minimize unwanted variability in longitudinal multisite studies of healthy elderly is crucial to validate probable DMN-based markers to predict the onset or track the staging of cognitive deterioration in elderly (Hedden et al., 2009; Vannini, 2012).

In this study group-ICA was used to extract DMN connectivity features and evaluate the TRT reliability of ICA-derived connectivity metrics, such as mean z-scores, coactivation cluster-sizes and coactivation spatial overlap across sessions, for ach rPNC method and each of the 13 sites in the consortium. Multisite evaluations of TRT reliability metrics were additionally investigated for each rPNC method. The following paragraphs report in part the research article published in Human Brain Mapping journal on March 2016 (Marchitelli et al., 2016).

3.2. Materials & Methods

3.2.1. MRI data acquisition

Fifteen clinical sites (13 MRI sites) participated in this study across Italy (Brescia, Verona, Genoa, Rome, Chieti, Perugia and Naples), Spain (Barcelona), France (Marseille, Lille, and Toulouse), Germany (Essen, Leipzig), Greece (Thessaloniki) and The Netherlands (Amsterdam) (Jovicich et al., 2016). The Brescia site was responsible for the coordination and analysis of the whole study and did not acquire MRI data. Each MRI site recruited 5 local volunteers within an age range of 50-80 years. Each subject underwent two MRI sessions completed at least 7 days (but no more than 60 days) apart at the site. This short period between the test and retest sessions was chosen to minimize biological changes that could affect the reliability of the measures and to mimic the variability expected from separate sessions, as measured in longitudinal studies. **Table**

3.1 summarizes information about age, gender and TRT interval times of the subjects scanned at each MRI acquisition site. All participants were volunteers with no history of major psychiatric, neurological or cognitive impairment (referred to as healthy in this study), were Caucasian and provided written informed consent according to the local ethical committee for each institution. Exclusion criteria were described in previous work (Jovicich et al., 2014).

The thirteen 3T MRI sites that participated in this study used a variety of MRI system vendors and models (Siemens, GE, and Philips). **Table 3.1** summarizes the main MRI system and EPI acquisition differences across sites. Only vendor-provided sequences were used. Each subject had a total of two resting state EPI acquisitions (RS-fMRI), one from the test session and one from the retest session. In each session the following single shot EPI acquisition parameters were common across sites: nominal voxel size 3 x 3 x 3 mm^3, TE = 30 ms, TR = 2.7 s, Ernst flip angle of 85^0, one-pass interleaved axial slices acquired (equidistant choice in Philips, default options in GE and Siemens) oriented parallel to the AC-PC line covering the full brain, 0.45 mm slice gap (15% slice thickness), 40 slices, 200 volumes, fat suppression, no parallel imaging. TR was chosen to be the shortest common possible value in the consortium allowing full brain coverage while keeping consistent temporal resolution across sites. The total resting state acquisition duration was 9 min, a duration that has been shown appropriate for reliable intrinsic connectivity results (Liao, 2013; Van Dijk et al., 2010; Zuo, 2013). Each MRI session included the acquisition of two T1 weighted anatomical scans (Jovicich et al., 2013). Subjects were instructed to keep their eyes closed, stay relaxed and try to avoid falling asleep.

Other acquisition parameters including head RF coil characteristics, pulse sequence and fat suppression methods were difficult to standardize due to system differences. The choices for these parameters were made based on the optimal or possible options available at the different platforms (**Table 3.1**). All images from multichannel coils were reconstructed online as the sum of the squares across channels. When allowed by the MRI system, images were reconstructed and saved disabling additional filtering options that could differ across scanners introducing different degrees of smoothing.

MRI site location		Subjects' age: mean±stdev, (range)	TRT time interval (days)	Gender (females/N)	3T MRI scanner	MR system software version	TX / RX coil		Fat suppression	In-plane matrix	tSNR Median (IQR)
Site 1	Verona	67.8±9.9 (26)	28 ± 23	2/5	Siemens	Allegra	VA25A	Birdcage	Fat Sat.	64x64	69 ± 8.8
Site 2	Barcelona	74.6±2.7 (6)	10 ± 3	5/5	Siemens	TrioTim	B17	Body 8-chan.	Fat Sat.	72x72	118.5 ± 8.1
Site 3	Leipzig	62.8±2.6 (6)	13 ± 3	3/5	Siemens	TrioTim	B17	Body 8-chan.	Fat Sat.	72x72	110 ± 18.3
Site 4	Marseille	66.0±8.3 (20)	23 ± 22	4/5	Siemens	Verio	B17	Body 12-chan.	Fat Sat.	72x72	134.5 ± 6
Site 5	Essen	52.4±1.5 (3)	11 ± 5	2/5	Siemens	Skyra	D11	Body 20-chan.	Fat Sat.	72x72	125 ± 62
Site 6	Naples	59.0±3.5 (9)	19 ± 15	2/5	Siemens	Biograph mMR	B18P	Body 12-chan.	Fat Sat.	72x72	94 ± 19.4
Site 7	Lille	64.2±5.3 (13)	15 ± 11	3/5	Philips	Achieva	3.2.2	Body 8-chan.	SPIR	80x80	120.5 ± 20
Site 8	Toulouse	59.2±4.5 (12)	14 ± 10	3/5	Philips	Achieva	3.2.2	Body 8-chan.	SPIR	80x80	100.4 ± 12.7
Site 9	Chieti	68.8±4.3 (11)	11 ± 5	5/5	Philips	Achieva	3.2.2	Body 8-chan.	SPIR	80x80	114.8 ± 51
Site 10	Perugia	60.8±10.3 (24)	10 ± 4	3/5	Philips	Achieva	3.2.2	Body 8-chan.	SPIR	80x80	118.6 ± 19
Site 11	Genoa	58.2±2.2 (5)	24 ± 17	2/5	GE	HDxt	15 M4A	Body 8-chan.	Water only	64x64	136.3 ± 19
Site 12	Thessaloniki	56.6±5.5 (5)	32 ± 8	3/5	GE	HDxt	15 M4A	Body 8-chan.	Water only	64x64	118.4 ± 20.3
Site 13	Amsterdam	62.8±8.2 (21)	11 ± 6	3/5	GE	HDxt	DV22.0 DV23.1	Body 8-chan.	Water only	64x64	158.6 ± 83.3

Table 3.1. Outline of MRI site, resting state fMRI acquisition and demographic differences.

[1]: A software upgrade took place during the study at site 13: DV22 was used for the first participant and DV23.1 for the rest.

3.2.2. Data preprocessing

Data preprocessing was performed in the individual space of each subject using a combination of FSL (Jenkinson et al., 2012) and AFNI (Cox, 1996) programs in the same order as they are listed. The following preprocessing steps were performed with NPC (**Figure 3.1**). The first 4 volumes were discarded (*fslroi*, FSL) to allow for steady state stabilization of the BOLD fMRI signal. Then, EPI volumes were realigned (*mcflirt*, FSL) and 6 head movement metrics were calculated; slice timing correction based on the slice acquisitions parameters at each site (*slicetimer*, FSL); non-brain voxels removal (*bet2*, FSL); temporal filtering with a bandpass filter (0.01 - 0.1 Hz) (*fslmaths*, FSL) to remove low and high-frequency signal fluctuations. Additional confounds, such as the 6 head movement metrics and their derivatives (*1d_tool.py*, AFNI), temporally filtered (*1dBandpass*, AFNI) as the main signal (Hallquist et al., 2013), were removed from the data using multiple linear regressions (*3dDeconvolve*, AFNI) for an overall of 12 regressors of no interest plus linear and quadratic trends. Neither regressors nor volumes were censored. 4D EPI volumes were spatially smoothed using 6 mm full-width at half-maximum Gaussian filter (*susan*, FSL) and normalized to mean signal intensity by a single factor (fslmaths, FSL).

Considering that variability in head-motion within and between subjects could confound our FC-fMRI estimates (Power et al., 2012; Van Dijk et al., 2012), head-motion was quantified using the framewise displacement (FD) timeseries derived from the 6 head movement metrics (Power et al., 2012). Mean and maxima FD were used to summarize head-motion in each subject and session.

3.2.3. Retrospective physiological denoising methods

Besides NPC, the following three methods for rPNC were evaluated: i) PESTICA correction (Beall & Lowe, 2007); ii) WM/CSF regression (Weissenbacher et al., 2009); and iii) FSL-FIX correction (Griffanti et al., 2014; Salimi-Khorshidi et al., 2014). Each of these methods followed a slightly different preprocessing workflow, as schematically shown in **Figure 3.1** and described below. All preprocessing pipelines, with or without rPNC, were performed in single subject space.

1. PESTICA (http://www.nitrc.org/projects/pestica/) is a freely available retrospective physiologic denoising algorithm (AFNI, MATLAB) that estimates the cardiac and breathing cycle directly from the BOLD fMRI data using temporal ICA (Beall & Lowe, 2007). The software defines 2 regressors-of-no-interest from the 2 most correlated independent components to spatial template references of heart beating and breathing, converts and regresses them out from the main signal using IRF-RETROICOR (Beall, 2010). Following the developers' recommendations, PESTICA v2.0 was run before slice timing correction. Default parameters suggested by the developers were used, such as temporal ICA dimensionality (15 components) and the full range for cardiac and respiratory estimators (48-85 bpm for cardiac and 10-24 bpm for breathing). All recommended PESTICA QA tools were used to verify that the algorithm performed properly: 1) visual inspection to assess the correct coregistration of the current EPI volume and the mean EPI template provided within PESTICA; 2) Plots of cardiac and breathing estimators; 3) coupling maps of cardiac and respiratory profiles.

2. Brain WM/CSF regression, here referred to as Tissue-based correction, is a common method to remove physiological confounds without a predefined model of noise (Birn et al., 2009; Bright & Murphy, 2013; Weissenbacher et al., 2009). This method relies on the assumption that BOLD fluctuations in WM and CSF are dominated by physiological noise, whereas BOLD fluctuations in GM combine signals from neural origin and physiological noise. One approach, the global regression method, removes the mean signal from the full brain. Therefore it might cause the loss of neural signal from GM. Instead, regression of the average WM and CSF time series from GM is expected to reduce the impact of physiological fluctuations in GM. Therefore the latter approach was chosen, here referred to as Tissue-based correction. Tissue-based correction is hence applied during multiple linear regressions together with movement confounds (**Figure 3.1**).

Tissue-based correction requires processing of T1 anatomical images to spatially define WM and CSF masks. In this study, for each subject, the first individual anatomical volume acquired at test was only used to define WM and CSF masks. Non-

brain tissue was initially removed from center-oriented T1 anatomical volumes (*bet*, FSL); the anatomical volume was then registered to the original EPI (*align_epi_anat.py*, AFNI); segmented in WM and CSF using partial volume segmentation (*fast*, FSL). Masks were thresholded at 0.99 (*3dcalc*, AFNI) to reduce the inclusion of gray matter voxels into the masks. To avoid gross coregistration and segmentation errors visual inspection was used. Two regressors of no interest were calculated from brain-extracted EPI as the mean time series within the non-gray matter masks (*3dcalc*, AFNI), temporally filtered as the main signal (*1dBandpass*, AFNI), and removed together with movement confounds by multiple linear regressions (*3dDeconvolve*, AFNI). The total number of nuisance regressors in this analysis pipeline was therefore 14 (6 head movement, 6 derivative of head movement, mean WM and CSF time series).

3. FSL-FIX (http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FIX) exploits single-subject spatial ICA to auto-classify noisy components and remove them from the main signal (Griffanti et al., 2014; Salimi-Khorshidi et al., 2014). Components classified as noise include physiological, head movement and other artifacts relative to image acquisition. The classification of noise components can be performed either by using one of several training files provided by the authors or by generating a data-specific training set on one's own data. In this study both approaches were investigated. Following the developers' suggestions, FSL-FIX was run after the NPC preprocessing pipeline. No changes were made to the preprocessing pipeline to be consistent across all rPNC methods examined. Temporal filtering was used consistently with common practice and works implementing FSL-FIX (Griffanti et al., 2014). Spatial smoothing was also performed, since it returned high reproducibility of the tSNR across multiple sites in our previous works (Jovicich et al., 2016). Single-subject ICA was run with automatic estimation of dimensionality using MELODIC on preprocessed data including the brain-extracted anatomical images (http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/MELODIC). From the various pre-defined FIX training sets offered, the one (herein referred to as FIX-standard) generated from a set of acquisition and preprocessing conditions as close as possible to those employed in this study was selected (Standard.RData: TR=3s, Resolution=3.5x3.5x3.5mm^3, session=6mins, default FEAT preprocessing including

default spatial smoothing).

The data-specific FIX training set (herein referred to as FIX-training) was generated separately for each MRI site using the two acquisition sessions of each subject (i.e., a total of 10 resting state runs per site). FIX-standard and FIX-training corrections were run using a signal/noise threshold of 75 and 20, respectively. Thresholds were determined as the minimum value to consistently remove at least those components that were visually related to physiological artifacts across subjects, sessions and sites. For FIX-standard this could not be achieved with a threshold lower than 75, whereas the implementation of site-specific classifiers in FIX-training allowed the usage of considerably smaller thresholds.

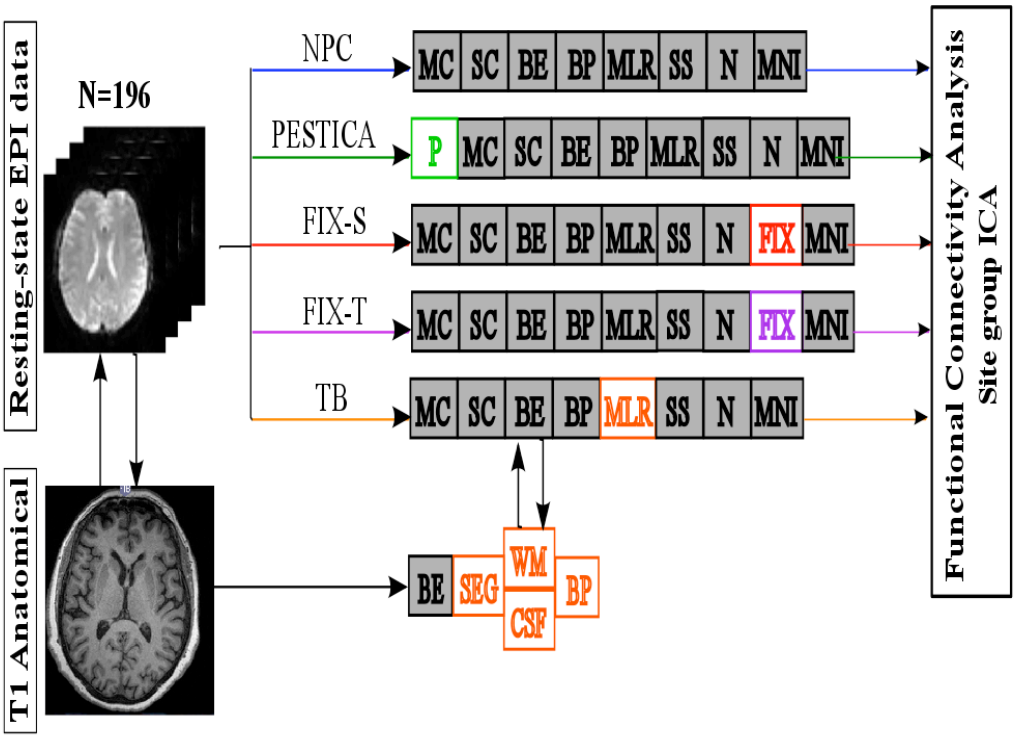

Figure 3.1. RS-fMRI preprocessing outline. The diagram shows main preprocessing steps for each rPNC method under investigation: NPC (blue); PESTICA (P, green); Tissue-based (TB, orange); FSL-FIX using standard (FIX-S, red) or site-specific classifier (FIX-T, violet). MC, motion correction; SC, slice-timing correction; BE, brain extraction; BP, band-pass filtering (0.001-0.01Hz); MLR, multiple linear regression; SS, spatial smoothing (FWHM = 6mm); N, mean intensity normalization; MNI, MNI coregistration.TB correction required anatomical image preprocessing: BE, brain extraction; SEG, segmentation; WM/CSF, white matter/cerebral spinal fluid regressors; BP, band-pass filtering of regressors (WM/CSF regressors were bandpass filtered). Anatomical BE images were also used for FSL-FIX.

3.2.4. DMN extraction methods

After preprocessing with each rPNC method in the individual space of each subject across all preprocessing workflows, group-ICA was performed using MELODIC (http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/MELODIC) to extract the site-specific group DMN, for each site and each rPNC method separately. For each site and rPNC method, all subjects and sessions (i.e. 10 resting state runs) were spatially normalized to the MNI template via linear (affine) registration (Jenkinson et al., 2002) and subsampled at a resolution of 4 mm isotropic voxels, then decomposed into 10 independent components (Jovicich et al., 2016), using the multisession temporal concatenation procedure in MELODIC. A higher number of components were not extracted to avoid splitting of the DMN (Abou-Elseoud et al., 2010; Jovicich et al., 2016).

An automatic selection procedure was used for each MRI site and rPNC method to select the group DMN from the 10 ICs in each dataset. In particular, an overlap measure was used to select the DMN using only the posterior regions (posterior cingulate and precuneus, left / right parietal cortex) from an independent functional DMN template (Rosazza et al., 2012) to avoid circularity. The component with the highest number of voxels in common with the template was chosen as representative of the DMN. Both our components and the DMN template were thresholded at z-scores > 2.3, $p < .01$ (Beckmann & Smith, 2004). The automatically selected group DMN for each site was visually inspected to assess overall resemblance with standard DMN patterns including its main coactivation areas.

For each site and rPNC method, dual-regression was then used to derive the single subject and session DMN from the site-specific group DMN (Beckmann, 2009; Zuo et al., 2010). Single-subject DMN volume maps were thresholded at $z > 2.3$, $p < 0.01$ (Beckmann & Smith, 2004). To account for spatial DMN variability across sessions and subjects, a functional cluster-criterion was used to define the DMN by four main clusters (*3dclust*, AFNI). These clusters were i) posterior cingulate and precuneus (BA31, BA30, BA29, BA23), ii) left/right parietal cortex (BA39, BA40, BA22, BA7) and iii) medial prefrontal cortex (BA9, BA10, BA32, BA24) (Franco et al., 2009).

To avoid inaccuracies for the definition of DMN activation maps, clusters were defined as made of voxels no more than 4mm apart and a cluster volume of at least 1800 micro-liters, per each ROI. These clusters were anatomically constrained by a reference DMN template derived from conducting a group-ICA overall the consortium with NPC implemented. For each rPNC method, spatial reproducibilityspatial reproducibility defined as the percent overlap between each individual DMN component and this reference template was calculated to quantify single-subject DMN variability across MRI sites after image processing.

Mean z-scores in the DMN (i.e., the mean z-score across all voxels in the four main clusters) and relative activation cluster-size (i.e. the total number of voxels in the four main nodes) were characterized for each subject, session, and rPNC method. Furthermore, the DMN activation cluster-overlap between sessions (i.e. the total number of common voxel coordinates between sessions) was also calculated for each subject and rPNC method.

3.2.5. Intra-site TRT reliability metrics

The main goal of this study was to evaluate the effects of different rPNC techniques on the precision of the DMN-derived measurements. To this end, a measure of TRT reliability within each site in the consortium was considered. Within-site TRT reliability was studied for each rPNC method using the following three metrics: 1) absolute percent change of mean DMN connectivity; 2) absolute percent change of DMN cluster-size, and 3) Jaccard index for DMN cluster-overlap across sessions (Bennett & Miller, 2010; Maitra, 2010; Meindl et al., 2010). These measurements were then used for the statistical analysis of within-site TRT reliability and inter-site reliability consistency across the consortium.

3.2.6. Inter-site reliability consistency metrics

For each DMN reliability measure and each rPNC condition, ICC(2,1) was used to quantify the degree of consistency of TRT reliability scores (here defined as inter-site reliability consistency) across all 13 sites in the consortium (Bennett & Miller, 2010). Considering subjects as “raters” and sites as “targets”, ICC measures the proportion of

variance between-sites out of the total variance. This formulation returns positive coefficients in the range [0,1] with values close to 1 indicating no strong site differences or biases in TRT reliability scores. Inter-site reliability consistency was measured using an in-house Matlab script; a leave-one-out approach was used to define the standard deviation of the ICCs.

3.2.7. Statistical measures

All statistical measures were performed using IBM SPSS Statistics for Macintosh, Version 22.0. Non-parametric Kruskal-Wallis tests was conducted to evaluate MRI site effects for each DMN-derived measurements and related TRT reliability scores under each rPNC approach. In the latter case, the variance estimated was used to calculate inter-site reliability consistency, for each DMN-derived measurement and rPNC method, respectively. The Kruskal-Wallis test was used on the various TRT reliability scores to obtain a statistical measure of site independence. The ICC analysis was used as a descriptive measure of inter-site reliability consistency across sites. Non-parametric Friedman test was conducted to evaluate rPNC effects in each site of the consortium and, if MRI site effects were not found (Kruskal-Wallis, $p > 0.05$), even on the pooled data across the consortium. rPNC method effects on the inter-site reliability consistency was also evaluated using Friedman test.

To increase the statistical power of the present analysis, along with the non-parametric analysis a 2-way ANOVA was also conducted on the pooled reliability scores across rPNC methods and sites for each DMN-derived measurement, separately. The indexes of kurtosis and skewness were examined to assess the distributional shape of TRT reliability scores and determine whether the assumption of normality was met (kurtosis range: [-2, 2] / skewness range: [-1, 1]) (Bulmer, 1979; Mat Roni, 2014).

Finally, bivariate Pearson's correlations were conducted to evaluate the existence of a relationship between the TRT reliability of connectivity measurements (z-scores, cluster-size, cluster-overlap) and across-session averaged movement estimates (mean FD). The significance level for all the statistics was set at $p < 0.05$. Statistics were corrected for multiple comparisons over all possible pairwise combinations (13 sites for Kruskal-Wallis and 5 rPNC methods for the Friedman test) using the method of Dunn-Bonferroni (Dunn, 1964) at $\alpha = 0.05$.

3.3. Results

3.3.1. Head-motion metrics

Overall low head movement was found: mean (SD) (0.11 ± 0.1 mm) with only 8% of time points exceeding 0.25 mm and 1% of them exceeding 0.5 mm. Mean FD range was (0.02 – 0.37 mm) where only the 4% of subjects exceeded higher mean FD than 0.25 mm. Maxima FD range was (0.07 – 3.32 mm) with roughly the 3% of subjects exceeding 1.5 mm (half-voxel size). Mean FD showed statistical differences across sites (Kruskal-Wallis, $\chi^2(12, n = 130) = 31$, $p = 0.002$ uncorrected) but not across sessions (paired t-test, p=0.99). Likewise, maxima FD showed site effects (Kruskal-Wallis, $\chi^2(12, n = 130) = 24$, $p = 0.02$ uncorrected) and no statistical difference between the two sessions (paired t-test, p=0.51). Given that the overall head movement was considered low and that ICA tends to separate movement components, we decided not to censor volumes in this study (Power et al, 2012). This also lets us use group-ICA on time series of the same length (196 vols).

3.3.2. Intrinsic DMN connectivity

Despite the low number of subjects (5 per sites, 2 repetitions for each), group-ICA successfully revealed DMN activation maps in all MRI sites and for all rPNC methods, consistently detected via automatic selection (**Figure 3.2**): selected DMN components substantially matched with DMN posterior regions for at least 50% in all MRI sites. Moreover, the difference between the first and second ranked components was very often higher than 30%, suggesting negligible variability between the selected DMN components across MRI sites (**Table 3.2**). Single-subject DMN maps were also successfully obtained using dual regression for each rPNC method and site (**Figure 3.3**).

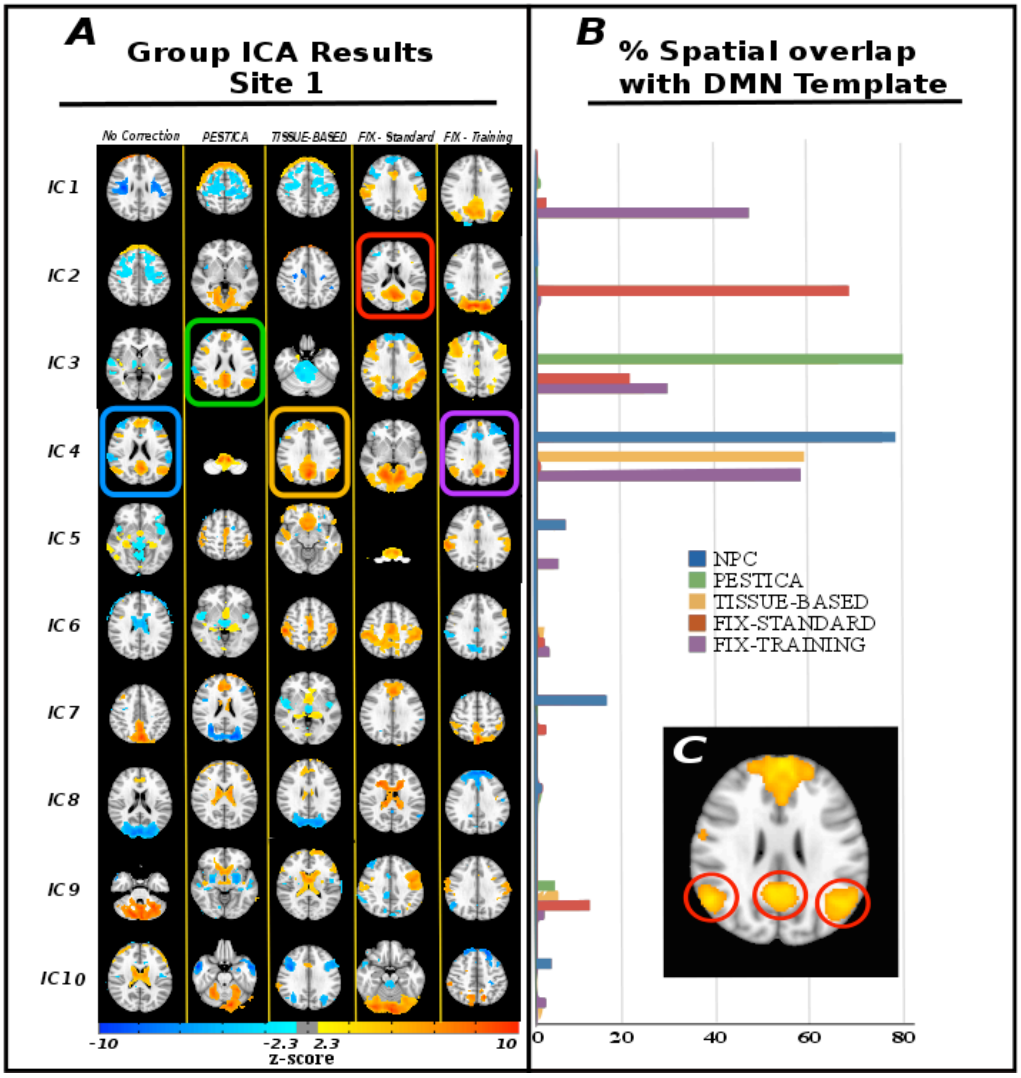

Figure 3.2. Automatic selection of the group DMN at a sample MRI site. (A) Group-ICA decomposition in 10 ICs (rows) for each physiological correction method (columns) is shown for one sample site (site 1). Activation maps include both positive (z > 2.3) and negative (z < -2.3) co-activations. Colored squares show the components that were automatically selected as the DMN in each rPNC condition. (B) Proportion of overlapping voxels between each independent component (IC) and the DMN template is reported for each rPNC method. Only positive z-score (z > 2.3) were considered for the selection. (C) The IC with maximum spatial overlap with the posterior regions of the DMN template was selected as the site-group DMN for each rPNC method respectively.

rPNC	Rank	Site 1	Site 2	Site 3	Site 4	Site 5	Site 6	Site 7	Site 8	Site 9	Site 10	Site 11	Site 12	Site 13
NPC	1st	71%	50%	75%	82%	69%	77%	71%	66%	59%	56%	71%	71%	71%
	2nd	14%	5%	43%	21%	22%	29%	36%	25%	26%	30%	27%	29%	63%
PESTICA	1st	75%	69%	69%	57%	67%	47%	62%	71%	50%	33%	59%	74%	54%
	2nd	4%	27%	4%	54%	26%	25%	17%	34%	15%	22%	13%	32%	13%
TISSUE BASED	1st	67%	59%	67%	72%	68%	76%	73%	66%	57%	65%	58%	74%	75%
	2nd	6%	34%	55%	37%	20%	28%	6%	11%	4%	22%	45%	34%	46%
FSL-FIX STANDARD	1st	64%	71%	77%	57%	68%	69%	64%	62%	62%	57%	57%	79%	76%
	2nd	19%	24%	47%	36%	52%	28%	49%	33%	13%	3%	46%	44%	28%
FSL-FIX TRAINING	1st	54%	69%	78%	82%	64%	78%	63%	66%	61%	65%	54%	66%	57%
	2nd	43%	52%	36%	37%	12%	39%	50%	36%	19%	16%	53%	41%	45%

Table 4.2. Spatial matching of two first-ranked group components with reference template. This table shows the percent overlap of the two first-ranked group independent components with the same reference template (all maps thresholded at $z > 2.3$) from each site and rPNC method, respectively. Overall, a spatial match of at least 50% was achieved in 63/65 cases for the selected DMN whereas only in 7/65 cases for the DMN competitor. A smaller difference than 30% between the two first-ranked components was found in 20/65 cases whereas a smaller difference than 15% in 10/65 cases. Although there is a certain degree of variability in percent overlap scores for the selected DMN across conditions, no systematic effects amenable to rPNC methods are observed.

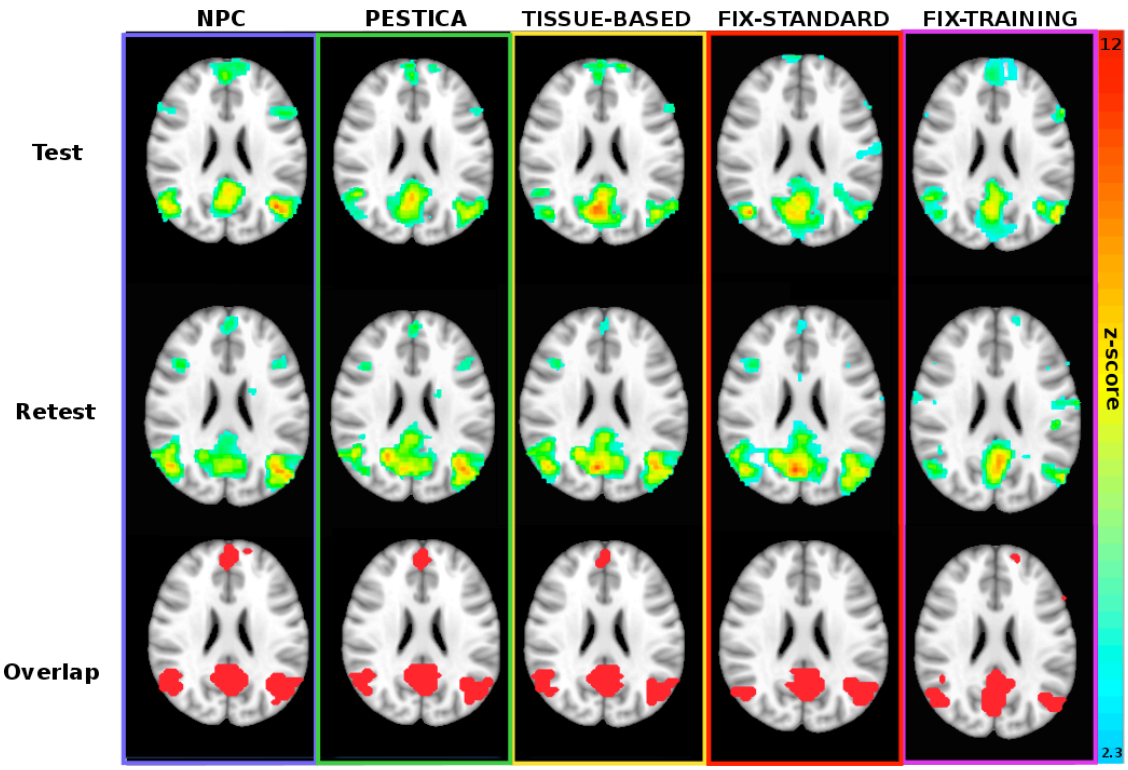

Figure 3.3. Single-subject DMN example. Reconstructed DMN map after dual regression on a sample single subject (site 1) is shown for each rPNC method (columns) at test (top row), retest (middle row) and its across-session spatial overlap (bottom row).

In NPC and FIX-training conditions, DMN patterns were either not found or did not exceed the threshold in one and five subjects, respectively. This allowed us to measure the median and its interquartile range (IQR) across subjects and sessions for the average DMN z-scores and cluster-size for each rPNC approach and MRI site (**Figure 3.4**). FIX-standard was found to classify 85% of single subject ICA components as noise across the entire consortium, whereas FIX-training, performed with a more conservative threshold, classified and removed 71%. FIX-standard removed much more components than FIX-training at each site except for site 2. In any case, both techniques were effective at attenuating physiological noise, overall. Spatial Reproducibility of single-subject DMN maps was overall medium-to-high within a range of 58-61%. Despite this convergence, a Kruskal-Wallis test revealed considerable MRI site effects in the spatial reproducibility of single-subject DMN maps. These were found for NPC, PESTICA, Tissue-based and FIX-standard and abolished only for FIX-training (Kruskal-Wallis, $\chi^2(12, n = 125) = 17$, $p = 0.2$).

A Kruskal-Wallis test also revealed significant MRI site effects in both DMN

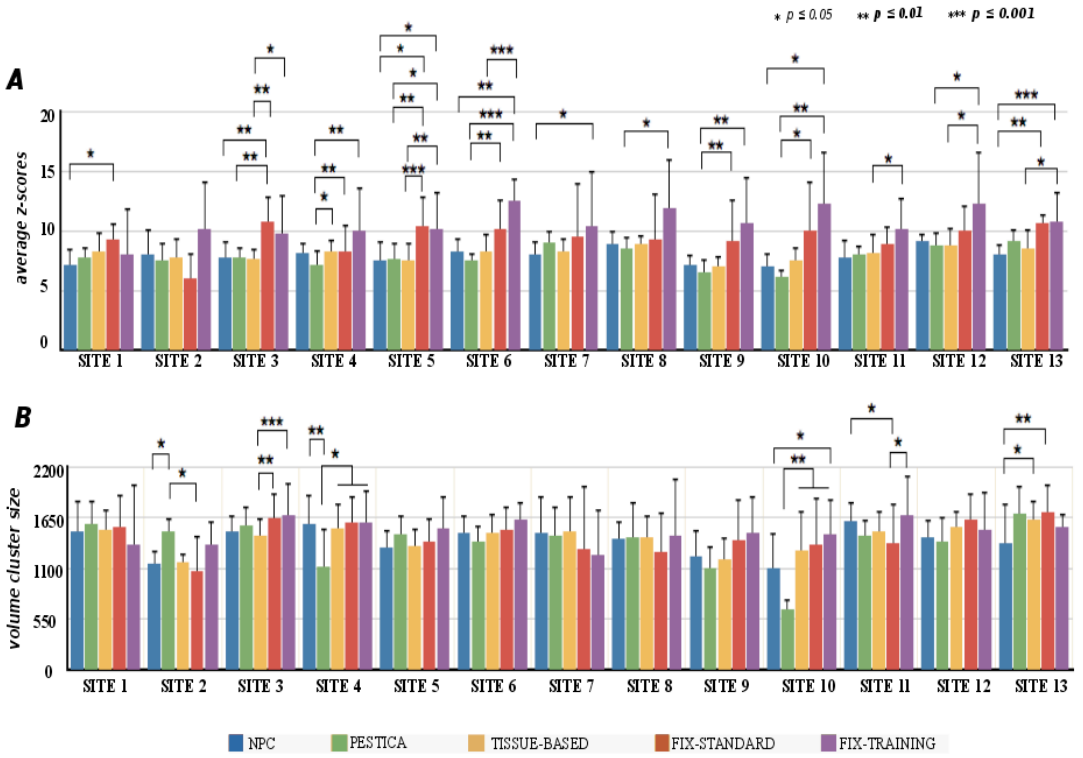

Figure 3.4. Within-Site intrinsic DMN connectivity and activation cluster-size results. Median (IQR) across subjects for supra-thresholded (A) mean z-scores (z > 2.3) and (B) volume cluster-size (N° active voxels at z > 2.3) within each site and for each PNC method. rPNC effects were evaluated within each site using Friedman test (p < .05, corrected with Dunn-Bonferroni method). Physiological correction using FIX tended to increase FC-fMRI but not cluster-size in the DMN across sites.

mean z-scores ($\chi^2(12, n = 129) = 29$; p=0.004 uncorrected) and cluster-size values ($\chi^2(12, n = 129) = 42$; p<0.001 uncorrected) when NPC is applied (**Table 3.3**). Similarly to single-subject DMN spatial reproducibility, MRI site effects persisted even after applying PESTICA, Tissue-based and FIX-standard methods. Only FIX-training canceled MRI site effects in DMN mean z-scores (Kruskal-Wallis, $\chi^2(12, n = 125) = 17$; p=0.2), cluster-size (Kruskal-Wallis, $\chi^2(12, n = 125) = 15$; p=0.3).

In general, FSL-FIX tended to increase median values of mean z-scores within the DMN nodes compared to the other methods: FIX-standard significantly increased median values of mean z-scores relative to NPC in 4 sites and FIX-training in 5 sites (**Table 3.4**). Nonetheless, FSL-FIX did not concurrently reduce activations in those brain regions prone to physiological noise such as WM and CSF in the selected DMN component (Andronache et al., 2013). The other rPNC methods did not exert any significant change on median values of mean z-scores relative to NPC in single sites.

Fewer rPNC effects were found for DMN cluster-size across sites relative to mean z-scores.

	CONNECTIVITY METRIC		TEST-RETEST RELIABILITY		
METHOD	**MEAN Z-SCORE**	**CLUSTER SIZE**	**MEAN Z-SCORE**	**CLUSTER SIZE**	**CLUSTER OVERLAP**
NPC	$\chi^2(12,129)=29$; $p=0.004$	$\chi^2(12,129)=42$; $p<0.001$	$\chi^2(12,64)=13$; $p=0.3$	$\chi^2(12,64)=19$; $p=0.08$	$\chi^2(12,64)=23$; $p=0.03$
PESTICA	$\chi^2(12,130)=54$; $p<0.001$	$\chi^2(12,130)=58$; $p<0.001$	$\chi^2(12,65)=7$; $p=0.9$	$\chi^2(12,65)=15$; $p=0.2$	$\chi^2(12,65)=29$; $p=0.004$
TISSUE BASED	$\chi^2(12,130)=28$; $p=0.006$	$\chi^2(12,130)=42$; $p<0.001$	$\chi^2(1265)=20$; $p=0.07$	$\chi^2(12,65)=13$; $p=0.3$	$\chi^2(12,65)=21$; $p=0.05$
FIX STANDARD	$\chi^2(12,130)=24$; $p=0.02$	$\chi^2(12,130)=30$; $p=0.003$	$\chi^2(12,65)=21$; $p=0.06$	$\chi^2(12,65)=17$; $p=0.1$	$\chi^2(12,65)=26$; $p=0.01$
FIX TRAINING	$\chi^2(12,125)=17$; $p=0.2$	$\chi^2(12,125)=15$; $p=0.3$	$\chi^2(12,60)=14$; $p=0.3$	$\chi^2(12,60)=8$; $p=0.8$	$\chi^2(12,60)=22$; $p=0.04$

Table 3.3. Statistical testing for site-effects. Kruskal-Wallis tests performed for each rPNC method (rows) and DMN-derived measurement (columns 1 - 2) and relative TRT reliability (columns 4 - 6) scores are here reported (p-values uncorrected). All DMN-derived measurements showed site-effects under NPC condition while FSL-FIX tended to reduce site effects; in particular FIX-training canceled site effects. In contrast, TRT reliability of intrinsic DMN connectivity and its activation cluster-size did not show site effects. Some subjects did not show DMN functional clusters at the given threshold in the NPC (1 subject) and FIX-training (5 subjects) conditions.

3.3.3. Intra-site TRT reliability metrics

Figure 3.5 shows the TRT reliability of the DMN connectivity metrics within each MRI site and rPNC method. We report the median (IQR) value across subjects at each site and rPNC method, for each DMN-derived measurement under investigation. No rPNC method significantly influenced TRT reliability relative to NPC at each single site for any DMN-derived measurement (Friedman test p>0.05, Bonferroni corrected).

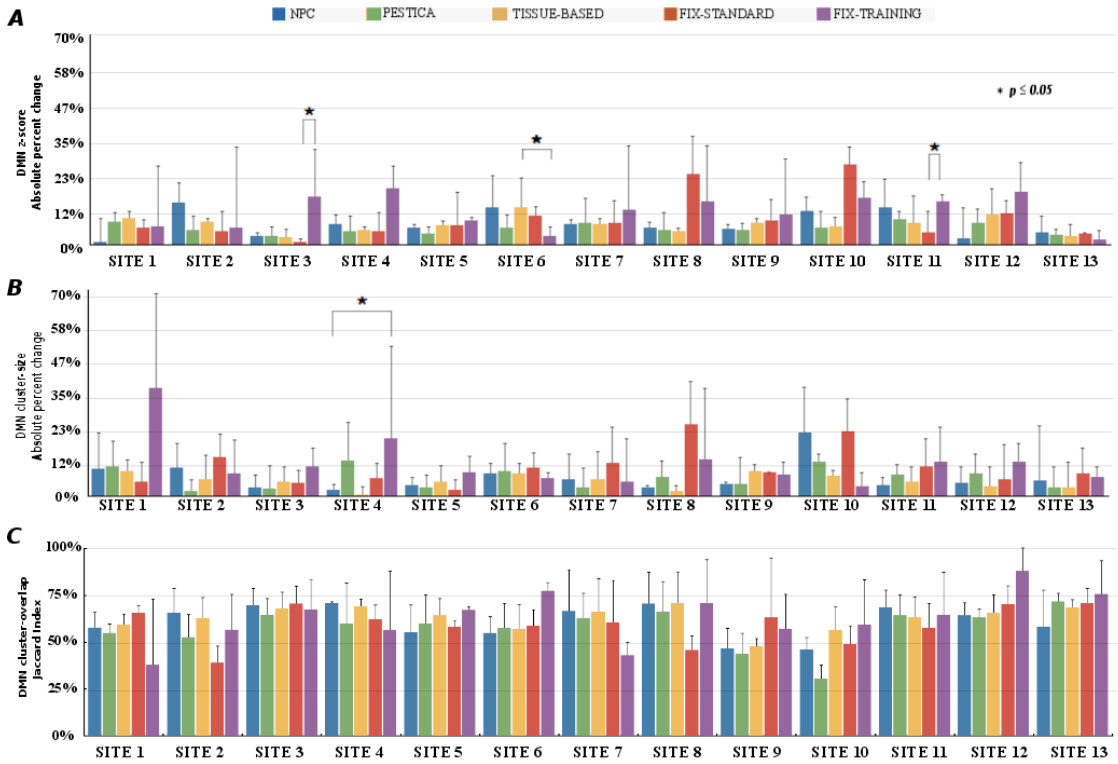

Figure 3.5. Within-site test-retest reliability of DMN-derived measurements for each rPNC method. Median (IQR) across subjects for TRT reliability scores in the DMN: (A) absolute percent change for FC-fMRI; (B) Cluster-size error (0% highest reliability, 100% lowest reliability); (C) Jaccard index of DMN spatial overlap (0% lowest reliability, 100% highest reliability). High TRT reliability was found for any rPNC method and DMN-derived measurement in all sites. rPNC effects were evaluated within each site using Friedman test ($p < .05$, corrected with Dunn-Bonferroni method). No systematic rPNC effects were found across the different sites. Statistically significant differences between rPNC methods were rarely found.

Furthermore, differences between rPNC methods were not systematic across MR sites. There were no statistically significant MRI site effects (Kruskal-Wallis, p>0.05) for the TRT reliability of mean z-scores in the DMN or the reliability of its

volume cluster-size under all rPNC method conditions. There were statistically significant MRI site effects (Kruskal-Wallis, $p < 0.05$) for the TRT reliability of cluster-overlap in the DMN. However, these effects did not survive multiple comparisons.

Therefore, all the TRT reliability scores of each rPNC method were pooled together across subjects and sites to examine overall TRT reliability (**Figure 3.6**). Absolute percent errors ranged from 5-11% for DMN z-scores and cluster-size reliability. DMN cluster-overlap was in the range 60-65%. There was a statistically significant difference in the TRT reliability of mean z-scores and volume cluster-size in the DMN, depending on the rPNC method implemented, mean z-scores: $\chi^2(4) = 14$; $p=0.009$, uncorrected; Cluster-Size: $\chi^2(4) = 21$; $p<0.001$, uncorrected; (**Table 3.4**). Multiple comparisons applied via pairwise comparisons revealed that the median (IQR) mean z-scores reliability error of PESTICA, 4.6% (6%), was statistically different from FIX-training, 10.6% (16.9%), ($Z=-3.6$, $p=0.003$). Pairwise comparisons revealed that the median (IQR) cluster-size reliability error of Tissue-based, 5.3% (6.3%) was statistically different from FIX-standard, 9.4% (12.1%), ($Z=-4.08$, $p<0.001$) and FIX-training, 7.3% (12.6%), ($Z=-3.03$, $p=0.03$). However, no rPNC method showed a statistically significant reduction of TRT reliability of mean z-scores or and cluster-size in relation to NPC within the DMN (**Figure 3.6**). Mean FDs did not correlate with absolute percent errors of DMN mean z-scores and cluster-size for all rPNC methods. However, the TRT reliability of cluster-overlap (i.e. the Jaccard index) was significantly anti-correlated with mean FD when using FSL-FIX (FIX-standard: $r(63) = -0.4$, $p< 0.01$ and FIX-training: $r(58) = -0.3$, $p < 0.01$).

Regarding the ANOVA framework, the distributions of TRT reliability scores were not normally distributed for mean z-scores (kurtosis: 2.5, skewness: 1.6) and cluster-size (kurtosis: 4.5, skewness: 1.9). However, a square root transformation was found to make these distributions more Gaussian-like (mean z-scores: kurtosis: -0.1, skewness: 0.5; cluster-size: kurtosis: 0.2, skewness: 0.6). The TRT reliability scores of the cluster-overlap were instead normally distributed (kurtosis: 0.3, skewness: -0.5). After this transformation, statistically significant rPNC method effects were found in

TRT reliability of mean z-scores ($F(4,12) = 4.5$, $p = 0.002$ uncorrected). In addition, statistically significant MRI site effects were found in the TRT reliability of both mean z-scores ($F(4,12) = 2$, $p = 0.02$ uncorrected) and cluster-size ($F(4,12) = 6$, $p < 0.001$, uncorrected). However, consistently with non-parametric statistical tests, both rPNC method and MRI site effects in the TRT reliability of both DMN-derived measurements did not survive multiple pairwise comparisons. Importantly, none of these metrics showed interaction effects.

Considering the TRT reliability of cluster-overlap (i.e. the Jaccard index), statistically significant MRI site effects ($F(4,12) = 8.3$, $p < 0.001$ uncorrected) and interaction between rPNC methods and sites ($F(4,12) = 1.9$, $p < 0.01$ uncorrected) were found. Multiple pairwise comparisons performed between MRI sites for each rPNC method revealed that MRI site effects were exacerbated under FIX-training correction and minimized under Tissue-based correction, respectively.

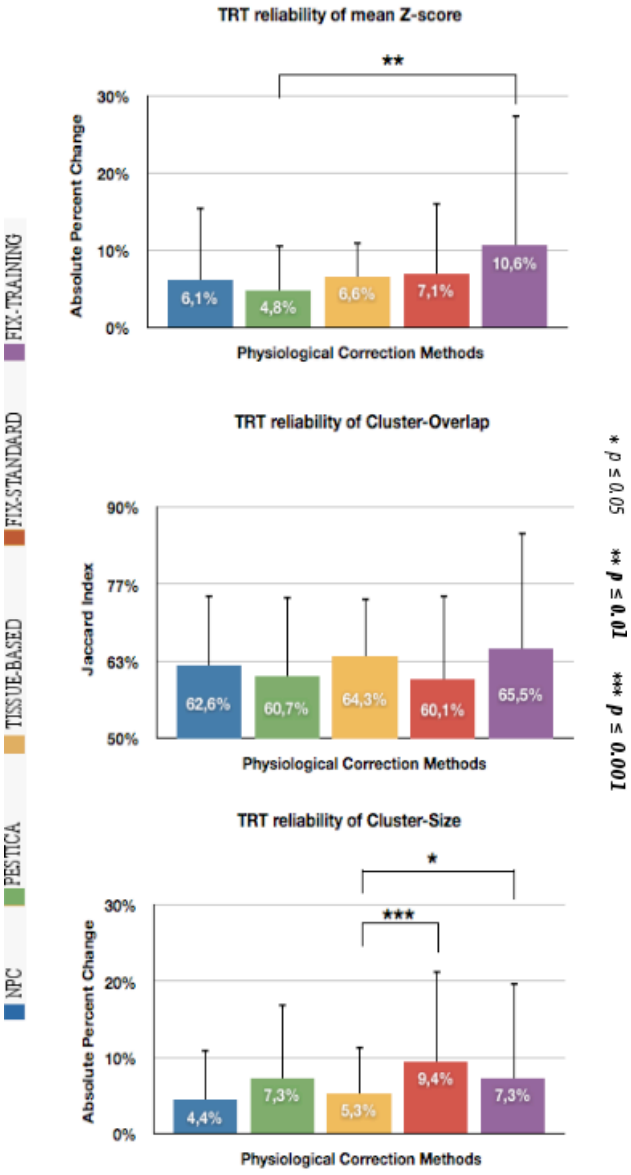

Figure 3.6. Overall test-retest reliability of DMN-derived measurements as function of rPNC method. Consortium median (IQR) of TRT reliability of each DMN-derived measurement: mean z-score (left), cluster-size volume (right), overlap-size (bottom). For each DMN-derived measurement, rPNC effects were evaluated using Friedman test ($p < .05$, corrected with Dunn-Bonferroni method). No rPNC method statistically reduced the TRT % error of FC-fMRI and cluster-size nor increased the Jaccard index for the cluster-overlap in the DMN.

	CONNECTIVITY METRIC		TEST-RETEST RELIABILITY		
	MEAN Z-SCORE	CLUSTER SIZE	MEAN Z-SCORE	CLUSTER SIZE	CLUSTER OVERLAP
SITE 1	$\chi^2(4)=13$; $p=0.1$	$\chi^2(4)=8$; $p=1$	$\chi^2(4)=2$; $p=1$	$\chi^2(4)=5$; $p=1$	$\chi^2(4)=6$; $p=1$
SITE 2	$\chi^2(4)=8$; $p=1$	$\chi^2(4)=14$; $p=0.1$	$\chi^2(4)=5$; $p=1$	$\chi^2(4)=6$; $p=1$	$\chi^2(4)=2$; $p=1$
SITE 3	$\chi^2(4)=24$; $p<0.01$	$\chi^2(4)=22$; $p<0.01$	$\chi^2(4)=11$; $p=0.4$	$\chi^2(4)=3$; $p=1$	$\chi^2(4)=6$; $p=1$
SITE 4	$\chi^2(4)=19$; $p=0.01$	$\chi^2(4)=17$; $p=0.03$	$\chi^2(4)=8$; $p=1$	$\chi^2(4)=16$; $p=0.04$	$\chi^2(4)=6$; $p=1$
SITE 5	$\chi^2(4)=31$; $p<0.01$	$\chi^2(4)=8$; $p=1$	$\chi^2(4)=3$; $p=1$	$\chi^2(4)=1$; $p=1$	$\chi^2(4)=1.6$; $p=1$
SITE 6	$\chi^2(4)=34$; $p<0.01$	$\chi^2(4)=8$; $p=1$	$\chi^2(4)=13$; $p=0.1$	$\chi^2(4)=31$; $p=1$	$\chi^2(4)=13$; $p=0.2$
SITE 7	$\chi^2(4)=10$; $p=0.5$	$\chi^2(4)=3$; $p=1$	$\chi^2(4)=2$; $p=1$	$\chi^2(4)=4$; $p=1$	$\chi^2(4)=5$; $p=1$
SITE 8	$\chi^2(4)=11$; $p=0.4$	$\chi^2(4)=1$; $p=1$	$\chi^2(4)=7$; $p=1$	$\chi^2(4)=8$; $p=1$	$\chi^2(4)=7$; $p=1$
SITE 9	$\chi^2(4)=21$; $p<0.01$	$\chi^2(4)=9$; $p=1$	$\chi^2(4)=2$; $p=1$	$\chi^2(4)=3$; $p=1$	$\chi^2(4)=4$; $p=1$
SITE 10	$\chi^2(4)=19$; $p=0.01$	$\chi^2(4)=29$; $p<0.01$	$\chi^2(4)=7$; $p=1$	$\chi^2(4)=9$; $p=0.9$	$\chi^2(4)=10$; $p=0.5$
SITE 11	$\chi^2(4)=14$; $p=0.1$	$\chi^2(4)=13$; $p=0.1$	$\chi^2(4)=10$; $p=0.4$	$\chi^2(4)=7$; $p=1$	$\chi^2(4)=5$; $p=1$
SITE 12	$\chi^2(4)=16$; $p=0.05$	$\chi^2(4)=5$; $p=1$	$\chi^2(4)=2$; $p=1$	$\chi^2(4)=5$; $p=1$	$\chi^2(4)=5$; $p=1$
SITE 13	$\chi^2(4)=26$; $p<0.01$	$\chi^2(4)=16$; $p=0.05$	$\chi^2(4)=4$; $p=1$	$\chi^2(4)=10$; $p=0.5$	$\chi^2(4)=6$; $p=1$
ALL SITES	-	-	$\chi^2(4)=14$; $p=0.1$	$\chi^2(4)=21$; $p<0.01$	$\chi^2(4)=7$; $p=1$

Table 3.4. Statistical testing for rPNC-effects. Friedman tests performed for each site (rows) and DMN-derived measurement (columns 1 - 2) and relative TRT reliability (columns 4 - 6) scores are here reported (adjusted p-values across 13 sites using Bonferroni correction). Mean z-score and volume cluster-size showed significant rPNC effects within many sites. In contrast, the TRT reliability of these DMN-derived measurements showed rPNC effects only rarely. The test was also performed on the scores from the entire consortium (last row) in case no site-effects were found for these estimates (see table 4.3).

3.3.4. Inter-site reliability consistency

Figure 3.7 shows inter-site reliability consistency (ICC) results for the three DMN-derived measurements (mean z-scores, volume cluster-size and cluster-overlap). A statistically significant difference was found in the inter-site reliability consistency of all these measurements (mean z-scores: $\chi^2(4) = 52$, $p < 0.001$; cluster-size: $\chi^2(4) = 54$, $p < 0.001$; cluster-overlap: $\chi^2(4) = 45$, $p < 0.001$), depending on the rPNC method implemented. For mean z-scores, pairwise comparisons revealed that the median (IQR) ICC values applying Tissue-based, 0.67 (0.04), and FIX-standard, 0.68 (0.04) was statistically different from NPC, 0.55 (0.02), ($Z = -3.5$, $p = 0.005$; $Z = -4.2$, $p < 0.001$; respectively). For cluster-size, pairwise comparisons revealed that median (IQR) ICC values are significantly reduced applying all rPNC methods excepted for FIX-standard (PESTICA, 0.47 (0.04), Tissue-based, 0.54 (0.02), and FIX-training, 0.28 (0.03)) compared to NPC, 0.65 (0.04), ($Z = 4.7$, $p < 0.001$; $Z = 3.2$, $p < 0.01$; $Z = 6.5$, $p < 0.001$, respectively). For cluster-overlap, pairwise comparisons showed that median ICC values for FIX-training, 0.61 (0.02), were significantly reduced compared to NPC, 0.74 (0.02), ($Z = 4.1$, $p < 0.001$).

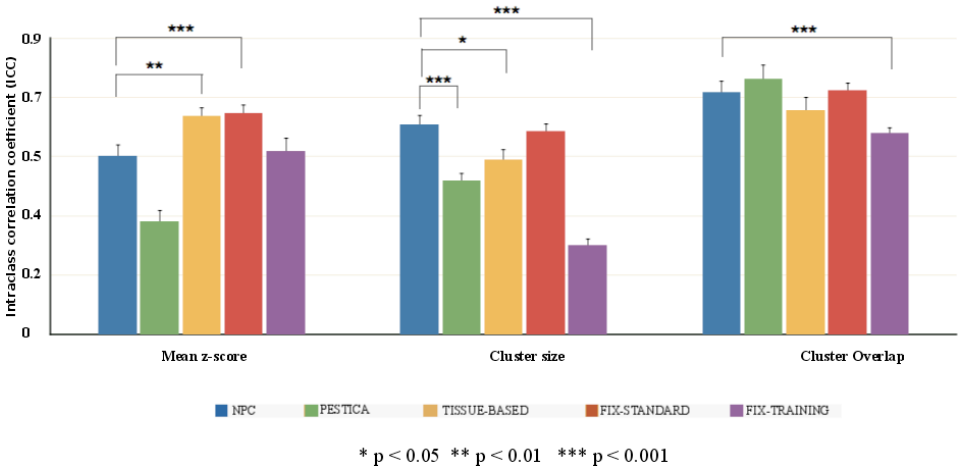

Figure 3.7. Inter-Site Reliability Consistency of DMN-derived measurements across sites. Inter-Site Reliability Consistency measured via ICC (mean ± standard deviation) is here shown for each DMN-derived measurement (bar groups) and rPNC method (color bars). For each DMN-derived measurement, rPNC effects were evaluated using Friedman test ($p < .05$, corrected with Dunn-Bonferroni method). Tissue-based and FIX-standard increase the inter-site reliability consistency of TRT reliability of average z-score reliability (left) with respect to NPC; in contrast, FIX-training increases variability of reliability estimates across different sites.

3.4. Limitations

General study design limitations were already discussed in recent morphometry (Jovicich et al., 2013) and diffusion (Jovicich et al., 2014) investigations from the Pharmacog consortium. This multisite study is limited in demographics for having investigated different individuals across sites and having acquired a rather small sample size at each site (five) reducing the statistical power in inter-site analyses. Moreover, TRT reliability analyses are limited by the acquisition of only two sessions, which could lead to variability understimation. Furthermore, multisite standardization came at the cost of reduced TR to image the entire brain, unavoidably affecting BOLD neural sensitivity and specificity, whereas shorter TR (< 1 s) acquisitions avoid aliasing of cardiac and respiratory fluctuations (Birn et al., 2008; Birn et al., 2006; Lund, 2001).

Besides the investigated rPNC methods, others including bayesian methods (Särkkä et al., 2012) or canonical correlations analysis (Churchill & Strother, 2013) could still be evaluated. From a methodological point of view, rPNC method effects were here ascertained only on model-free ICA-based DMN connectivity but graph-based measurements of network architecture are currently receiving wide attention in the context of monitoring physiological and pathological brain changes (Friston, 2011; Sporns, 2014; Stam, 2014); rPNC method effects should therefore be evaluated in future TRT reliability graph-based FC-fMRI studies.

4

Experiment 2: Influence of slice-order acquisition and head-motion correction methods on the DMN reliability in healthy adults

4.1. Introduction

The second challenge addressed in the present thesis relates to head-motion in longitudinal RS-fMRI studies of the DMN. In this 4T resting-state study of healthy young adults, we evaluated whether different combinations of ascending slice-order protocols, namely interleaved or sequential, and head-motion correction methods, namely volume-based or slice-based, would influence the TRT reliability of intrinsic DMN connectivity. Our initial hypothesis was that, in single-shot multislice echo planar imaging, sequential slice-order acquisitions could minimize BOLD sensitivity to in-scanner head-motion compared with interleaved acquisitions and hence lead to more reliable DMN connectivity metrics in longitudinal resting-state connectivity studies (Cheng & Puce, 2014). To this extent, recent retrospective methods providing for both in-volume and inter-volume motion correction could be advantageous over standard rigid-body volumetric methods for both slice-order acquisition sequences (Beall & Lowe, 2014).

As previously described in this thesis, the in-scanner head-motion confounding factor cannot be easily isolated during MR scanning sessions for being very subject-specific and beyond the experimenter control (Van Dijk et al., 2012). This biological artifact exerts a global effect even in exiguous amount (< 1 mm) causing alterations in

FC-fMRI patterns, leading to an increase in spatial short-distance connections and reduction in long-distance and anterior-posterior connectivity (Murphy et al., 2013).

These effects introduce unwanted variability across subjects making it challenging to compare different groups of individuals characterized by different head-motion profiles (Power et al., 2012; Satterthwaite et al., 2012; Van Dijk et al., 2012; Zeng et al., 2014). Moreover, head-motion is associated with reliability loss in longitudinal resting-state studies (Guo et al., 2012). Therefore, it is important to evaluate how specific approaches aimed at minimizing head-motion can influence the TRT reliability in longitudinal studies.

While the most commonly adopted resting-state acquisition protocols consider interleaved slice-order sequences to minimize cross-talking effects, this choice might be sub-optimal to account for head-motion and prone to spin-history artifacts. In contrast, sequential slice-order sequences would represent a diametrically opposite technical solution with overall reduced sensitivity to head-motion (Kim et al., 2008; Sladky et al., 2011).

When it comes to data preprocessing, commonly implemented volumetric coregistration methods (a.k.a. motion-correction or volume-realignment) such as FSL-FLIRT or AFNI-3dvolreg are also suboptimal for being developed withouth considering that each excitation pulse is targeted to one slice at a time and that the head is not still but moving across all slices (Murphy et al., 2013). This means that these software do not account for motion that occur on a scale of less than a TR (in-volume) because they are designed to deal only with rigid-body motion occurring across volumes (inter-volume).

Other jointly applied approaches such as multiple linear regressions of the estimated rigid-body head translations and rotations (Satterthwaite et al., 2013), “censoring” highly motion-compromised volumes (FD > 0.25 mm) (Power et al., 2012), ICA-based techniques to isolate motion-related spatio-temporal characteristics (Griffanti et al., 2014; Pruim et al., 2015; Schopf et al., 2010) can further reduce non-linear motion-related variability but cannot account for the consequences of head-motion on activation timing and patterns of excitation that occur across slices during EPI volume acquisition.

To the best of our knowledge, no resting-state TRT reliability studies have previously investigated how combiantions of these acquisition and motion correction methods would influence the stability of intrinsic DMN connectivity metrics. This 4T TRT reliability study evaluates whether some combinations of slice-order protocols and head-motion correction methods would improve the stability of ICA-based FC-fMRI measures within the default-mode system of healthy young adults. FC-fMRI was measured in the entire DMN and its main four regions, separately. For each ROI, mean z-scores and cluster-sizes were collected as measures of FC-fMRI. TRT reliability analysis included examination of the absolute percent change of mean z-scores, ICC evaluations of mean z-scores and inter-session spatial reproducibility using the jaccard index.

4.2. Materials & Methods

4.2.1. MRI data acquisition

A 4T Bruker Medspec scanner (Bruker Medical, Ettlingen, Germany) with a birdcage-transmit and an 8-channel receive head coil (USA Instruments, Inc., Ohio) was used to acquire resting-state scans from 24 healthy young volunteers (12 female, 27.0 ± 5.3 y). Subjects were scanned in two sessions, in average 4.2 ± 4.4 days apart. During each RS-fMRI session subjects were asked to relax, keep their eyes closed, stay still avoiding falling asleep or engaging in structured thoughts. In each session, one 3D T1-weighted MP-RAGE (TR/TE 2,700/4.18 ms; 1 mm^3 voxel), and two resting state fMRI (rsfMRI) runs (TR/TE 2.2s/30 ms, with 37 AC-PC parallel slices, voxel size 3 mm^3, slice-gap 0.6 mm, 200 volumes) were acquired. The RS-fMRI scans were acquired once with ascending interleaved and once with ascending sequential 2D slice-order protocols. Data from 1 participant were excluded from the analysis for being collected with a smaller number of TRs.

4.2.2. Head-motion correction methods

Below the two head-motion correction methods applied and compared in this study are outlined: volume-based head-motion correction, here referred to as VOMOCO

(Jenkinson et al., 2002), and slice-based head-motion correction, here referred to as SLOMOCO (Beall & Lowe, 2014).

VOMOCO was performed using a rigid-body motion correction method (http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/MCFLIRT) based on the affine registration tool in FSL (FLIRT). For each subject and session, the train of 3D brain volumes (N=194) was registered to the middle reference volume (97th) using linear sinc interpolation with 6 degrees of freedom (DOF). This procedure yields a transformation matrix of 6 volumetric movement displacement parameters.

SLOMOCO was performed using SLice-Oriented MOtion COrrection (http://www.nitrc.org/frs/?group_id=361). SLOMOCO is a publicly available slice-based motion regression method that corrects each slice for movement within and between volumes using AFNI programs. During within-volume motion correction each slice undergoes an in-plane co-registration. Motion estimates are regressed from each voxel including movement information from adjacent slices. Thereafter, between-volume motion correction is performed for each slice of interest (all other slices are replaced with their mean timeseries across time). This procedure yields a 6 DOF transformation matrix re-assembling the 3 DOF movement time-series resulting from both correction steps and matching the average of their variances, respectively. Of note, time-series from between-volume correction are variance-normalized and all slices have the same variance thereof (Beall & Lowe, 2014).

Both motion-correction methods produced 6D head-motion timeseries associated with 3D head translation and 3D head rotation in each subject and session and under each slice-order protocol. These metrics were simplified to a 1D timeseries to quantify the amount of head motion: 3D head rotation timeseries were converted from radians to millimeters in terms of the corresponding displacement on a r=50 mm sphere, which represents the average distance between the cortex and head center (Siegel et al., 2014); then, the straight sum across timeseries (in absolute value) was calculated to obtain 1D volume-to-volume head-displacements. The median value of the 1D timeseries was finally calculated to summarize head-motion in a scalar.

4.2.3. Data preprocessing

The preprocessing of the functional EPI data was performed in the individual space of each subject using a combination of FSL (Jenkinson et al., 2012) and AFNI (Cox, 1996) programs (**Figure 4.1**). Non-brain voxels were removed from the EPI volume (bet2, FSL) and, to remove high-frequency signal fluctuations, a temporal filtering with a bandpass filter (0.01 - 0.1Hz) was used on both the EPI volume (fslmaths, FSL) and the 6 movement metrics (1dBandpass, AFNI) (Hallquist et al., 2013). Multiple linear regressions (3dDeconvolve, AFNI) were used to remove the 6 movement metrics and their relative derivatives (1d_tool.py, AFNI) plus second-order polynomials from the main signal.

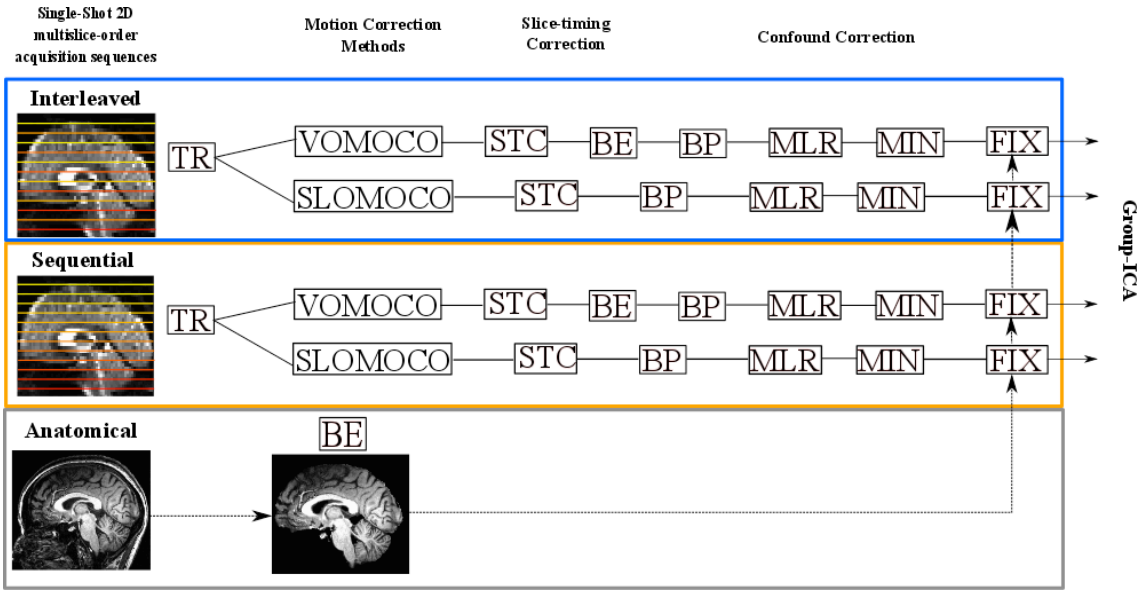

Figure 4.1. Pre-processing workflow. The figure illustrates the preprocessing pipeline under each slice-order acquisition condition, either interleaved (blue rectangle) or sequential (orange rectangle) and movement correction method, either volume-based (VOMOCO) or slice-based (SLOMOCO). TR, denotes the removal of the first 6 TRs from the BOLD timeseries; STC, slice-timing correction (interleaved / sequential); BE, brain extraction (included in SLOMOCO); BP = band-pass filtering (0.01–0.1 Hz); MLR, multiple linear regression (12 head movement timeseries (6 head movement timeseries + 6 derivatives) filtered as BP; MIN, mean intensity normalization; FIX, FSL-FIX implementation. The bottom row indicates the anatomical brain extraction step needed for use in FSL-FIX.

Individual brain volumes were normalized to mean signal intensity by a single factor (fslmaths, FSL) and registered to brain-extracted anatomical volumes using Boundary-Based-Registration (Greve & Fischl, 2009) and decomposed into 25 spatial independent components using single-subject spatial ICA algorithms implemented in MELODIC (http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/MELODIC). FSL-FIX (Griffanti et al., 2014; Salimi-Khorshidi et al., 2014) was implemented to detect and remove physiological noise and MRI hardware-related artifactual components using a standard classifier (Standard.RData: TR=3s, Resolution=3.5x3.5x3.5mm3, session=6mins, default FEAT preprocessing) and a medium classification threshold (50/100) as a criterion to distinguish between signal and noise associated components. To remove physiological noise confounds FSL-FIX uses the non-gray matter tissue masks (WM & CSF) and the 6 motion parameters obtained after head-motion correction. The tissue masks are derived from the brain extracted anatomical volume, which was obtained using OptiBET for each single subject and session (Lutkenhoff et al., 2014).

Although the contribution of head motion is substantial at 4T field strength (Hutton et al., 2011; Triantafyllou et al., 2005) and spatial smoothing can counteract motion-induced BOLD modulations (Triantafyllou et al., 2006), no spatial smoothing was applied since its application is discouraged in conjunction with FSL-FIX (Salimi-Khorshidi et al., 2014) and does not improve the detectability of the DMN (Molloy et al., 2014). Moreover, given that ICA tends to separate motion components, we decided not to censor volumes in this study (Power et al., 2012). This also lets us use FSL-FIX and perform group analysis on time series of the same length (194 vols).

4.2.4. DMN extraction methods

After preprocessing, group-ICA was conducted across all subjects and sessions for each slice-order protocol and head-motion correction method (i.e. 48 resting state runs per condition), using MELODIC to extract the group DMN (**Figure 4.2**). All preprocessed images were spatially normalized to the MNI template via application of linear (affine) FLIRT (Jenkinson & Smith, 2001) and subsampled at a resolution of 3 mm isotropic voxels, before decomposition into 15 independent components, using the multi-session temporal concatenation procedure in MELODIC. A higher number of

components were not extracted to avoid splitting of the DMN (Abou Elseoud et al., 2011; Jovicich et al., 2016).

For each condition, the DMN was visually selected among the 15 components as the one that included the most common DMN nodes: medial prefrontal and anterior cingulate cortex, posterior cingulate cortex and precuneus, left/right parietal corteces, using a threshold of z-score > 2.3 (Franco et al., 2009). A reference template (Rosazza et al., 2012) was also used to qualitatively verify the degree of overlap of the group component map.

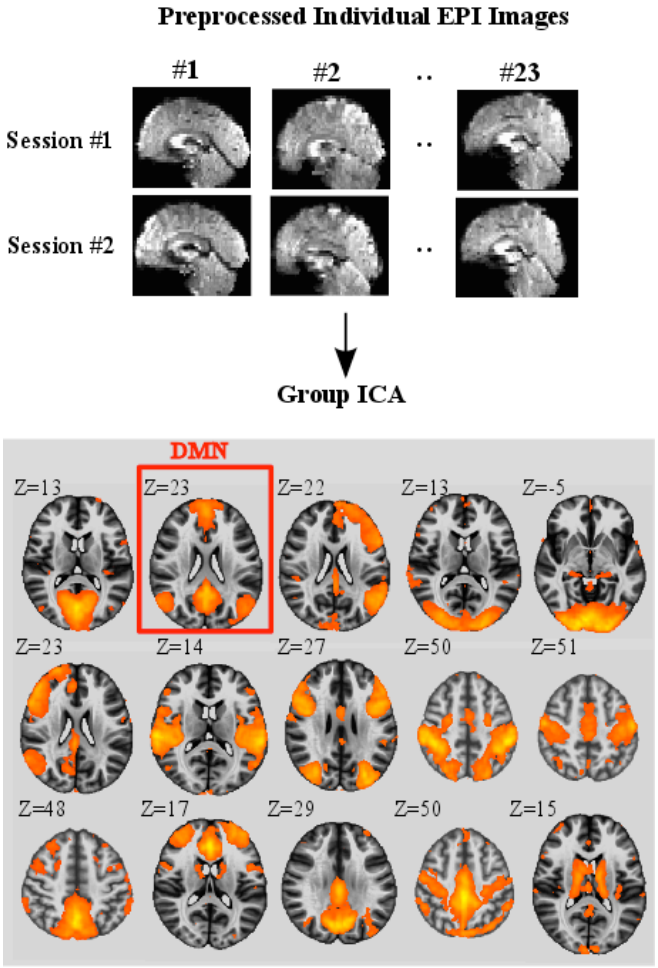

Figure 4.2. Group-ICA results at 15 IC dimensionality. The figure illustrates a sample dataset of preprocessed EPI volumes (Interleaved + SLOMOCO) for each MRI session (session 1 or test, session 2 or retest) (top) and following group ICA using 15 components on the group of 48 sessions (bottom). Statistical maps were thresholded at z > 2.3. The second component in the top-left corner was visually selected as the DMN in this example.

For each slice-order protocol and head-motion correction method, dual-regression was then used to derive the DMN for each single subject and session **(Figure 4.3)** from the selected group component (Beckmann et al., 2009). Individual DMN volume maps were thresholded at $z > 2.3$, $p < 0.01$ (Beckmann & Smith, 2004). To account for spatial DMN variability across sessions and subjects a cluster analysis was run to detect the four main functional co-activated clusters that characterize the DMN (*3dclust*, AFNI). These clusters are the posterior cingulate and precuneus (including BA31, BA30, BA29, BA23), the left/right parietal cortex (including BA39, BA40, BA22, BA7) and prefrontal and anterior cingulate areas (including BA9, BA10, BA32, BA24) (Franco et

al., 2009). To avoid inaccuracies for the definition of DMN activation maps, clusters were defined as made of voxels no more than 4mm apart and a cluster volume of at least 1800 micro-liters in size, per each ROI. These clusters were anatomically constrained by a reference DMN template (Rosazza et al., 2012). Visual inspection was done to verify that clusters overlapped with the main regions attributed to the DMN.

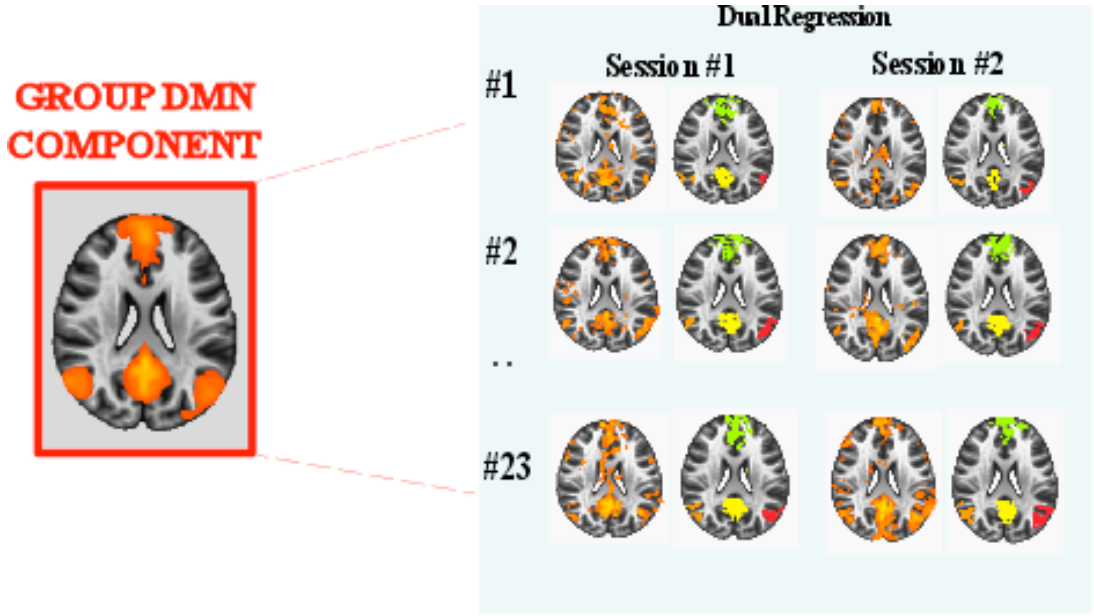

Figure 4.3. Dual Regression Analysis. The figure illustrates dual-regression results relative to the DMN from the same sample subjects shown in Figure 13. All statistical maps are thresholded at $z > 2.3$. In the right panel, individual DMN activations before (left) and after the cluster-analysis (right) are shown for some subjects for each session. Colors indicate ROIs (PCC: yellow; ACC: green; LPC: red/orange; DMN: all ROIs). FC-fMRI metrics derived from these maps were forwarded to the TRT reliability analysis.

Mean z-scores in the DMN (i.e., the mean z-score across all voxels in the four main clusters) and relative activation cluster-size (i.e. the total number of voxels in the four main nodes) were characterized for each subject, session and condition. In addition, mean z-scores and relative cluster-sizes were calculated for the posterior cingulate, medial prefrontal cortex and parietal cortex, separately.

4.2.5. TRT reliability metrics

The main goal of this study was to evaluate the effects of different slice-order acquisition strategies and head-motion correction techniques on the longitudinal precision of intrinsic DMN connectivity metrics. For each of these condition and for

each ROI under consideration, three TRT reliability metrics were here calculated on suprathresholded connectivity maps ($z > 2.3$): 1) the absolute percent change of mean ROI z-scores to measure the inter-session error in mean z-scores for each subject (Bennett, & Miller, 2010); 2) the intraclass correlation coefficient ICC(2,1) and relative confidence intervals (CI) (McGraw & Wong, 1996) of mean ROI z-scores to measure the proportion of inter-session variance out of the total variance in the entire sample (Shrout & Fleiss, 1979) and 3) the Jaccard index, to measure voxel-wise spatial inter-session convergency of cluster activations in each subject (Jovicich et al., 2016).

4.2.6. Statistical evaluations

Statistical evaluations were performed in IBM SPSS Statistics for Macintosh, Version 22.0. A paired t-test was used to evaluate fatigue effects in the tSNR across slice-order protocols. Bivariate Pearson's correlations were computed to evaluate the existence of relationships between FC-fMRI and head movement measures in all the ROIs under investigation and for each combination of slice-order protocol and head-motion correction method. In the specific case of TRT reliability examinations, volume-to-volume head movement parameters were averaged across sessions.

The Friedman test was conducted to evaluate both slice-order acquisition and head-motion correction method effects in head-motion derived metrics (i.e. median volume-to-volume displacement) across all subjects and sessions (N=47). The same non-parametric test was also used to assess the same effects in mean z-scores and in both TRT reliability measures, for the entire DMN and its single ROIs, respectively.

Statistical significance level was set to $p < .05$. Statistics were corrected for multiple comparisons over all possible pairwise combinations (2 slice-order x 2 movement correction methods) using the method of Dunn-Bonferroni (Dunn, 1964) at $\alpha = 0.05$.

4.3. Results

4.3.1. tSNR & head-motion metrics

No significant difference was found in tSNR measures for interleaved (22.2 ± 4.3) and sequential (22.6 ± 4.4) slice-order protocols; t(46) = -4.9, p = 0.63. Overall estimated head movement parameters from VOMOCO and SLOMOCO are summarized in **Table 4.1**. The Friedman test performed across all slice-order protocols and head-motion correction methods revealed statistical significant differences in the amount of estimated motion (Friedman, $\chi^2(3, n = 47) = 112.9$, $p < .001$ uncorrected). Pairwise comparisons revealed that median volume-to-volume displacements were statistically different between head-motion correction methods, (Interleaved: Z = -2, $p < 0.001$; Sequential: Z =-2, $p < 0.001$) and not statistically significant between slice-order protocols, (VOMOCO: Z=.06, p=.81; SLOMOCO: Z=.02, p=.94).

In summary, in this group of subjects, the amount of movement estimated from both head-motion correction methods was relatively low compared to the voxel size. The movement estimates were not significantly affected by slice-order sequences but they were significantly affected by the head-motion correction method used, with SLOMOCO giving larger movement estimates.

	TSNR & HEAD-MOTION ESTIMATES			
	INTERLEAVED		**SEQUENTIAL**	
TSNR	22 ± 4		23 ± 4	
	VOMOCO	**SLOMOCO**	**VOMOCO**	**SLOMOCO**
MEDIAN	0.19 ± 0.17 mm	1.1 ± 0.9 mm	0.2 ± 0.1 mm	1.2 ± 1.1 mm
MAXIMA	0.8 ± 0.7 mm	2 ± 1 mm	1.2 ± 1.4 mm	2.8 ± 2.6 mm

Table 4.1. Head movement statistics after head motion correction. The table shows averaged median and maxima head movement displacements across individuals for each slice-order acquisition sequence and head movement correction method. Bottom line reports the tSNR prior any head movement correction. Head movement displacement timeseries were derived from the straight sum of the absolute value of 6 head movement parameters of 194 time points in length. Head movement estimates and the tSNR are similar across slice-order acquisition sequences but differ across head movement correction methods.

4.3.2. Intrinsic DMN connectivity

Notwithstanding differences in movement estimates across head-motion correction methods during the pre-processing of the data, group-ICA revealed DMN co-activation patterns across all combinations of slice-order protocols and head-motion correction method. No subject was discarded during data post-processing since dual-regression methodology revealed individual DMN maps across all subjects, sessions and combinations-of-inteerst, and the cluster analysis returned intact DMN maps which allowed measurability of mean z-scores in all ROIs (**Figure 4.4)**. Descriptive and inferential statistical results are reported in **Table 4.2.** As for the full DMN, significant statistical differences in mean z-scores across slice-order protocols and head-motion correction methods were found ($p < 0.001$ uncorrected).

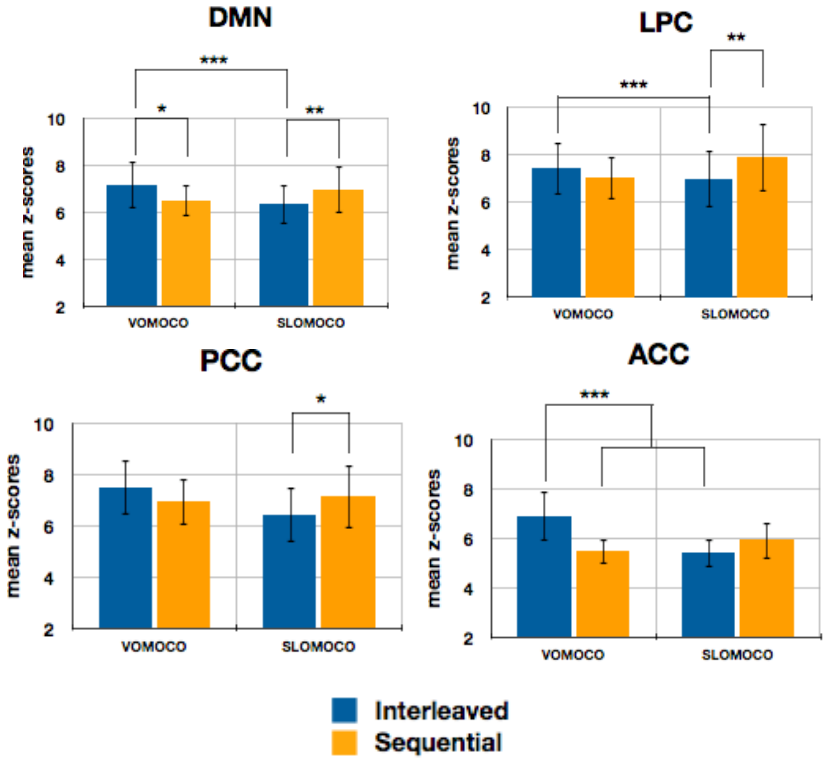

Figure 4.4. Intrinsic FC-fMRI in the DMN and its main nodes. Mean z-scores averaged across subjects (error bars are the standard deviation between subjects) in four ROIs (from left to right): the entire DMN (DMN), the posterior cingulate cortex (PCC), lateroparietal lobes (LPC), and anterior cingulate (ACC). For each ROI, results are differentiated for movement correction method, within which slice-order conditions are displayed (Interleaved, blue; Sequential, orange). Mean z-scores are sensitive to slice-order acquisitions and movement correction methods in all ROIs (corrected).

Pairwise comparisons revealed that mean z-scores in the entire DMN were statistically different between head-motion correction methods, (Interleaved: Z = 1.3, p < 0.001) and statistically different between slice-order protocols, (VOMOCO: Z = 0.08, p = 0.02; SLOMOCO: Z = - 0.09, p = 0.005). Statistical significant effects were also found in each single node of the DMN (PCC: p = 0.025 uncorrected; ACC: p < 0.001 uncorrected; LPC: = p < 0.001 uncorrected).

	MEAN Z-SCORES				Statistics
	VOMOCO		SLOMOCO		
	INTERLEAVED	SEQUENTIAL	INTERLEAVED	SEQUENTIAL	
DMN	7.3 ± 0.9	6.6 ± 0.7	6.5 ± 0.8	7 ± 1	Friedman X^2 (3, n=47)=26.18 p < .001
PCC	7.4 ± 1.1	7.2 ± 0.9	7.1 ± 1.2	8 ± 1.4	Friedman X^2 (3, n=47)=9.36 p < .025
ACC	7 ± 1	5.6 ± 0.5	5.6 ± 0.5	6.1 ± 0.7	Friedman X^2 (3, n=47)=56.03 p < .001
LPC	7.6 ± 1.1	7 ± 0.9	6.5 ± 1	7.2 ± 1.2	Friedman X^2 (3, n=47)=23.7 p < .001

Table 4.2. FC-fMRI results. Average (SD) mean z-score across participants under each condition (columns) for each ROI (rows). The last column provides non-parametric analysis of variance output (uncorrected). Significant statistical effects were found across conditions in each single ROI ($p < 0.05$). Except for PCC (slice order effects only), for all other areas the effects were significant for both slice order and motion correction method.

In the PCC, pairwise comparisons revealed that mean z-scores were statistically different between slice-order protocols (SLOMOCO: Z = -.76, p = .03) but not across head-motion correction methods. In the ACC, pairwise comparisons revealed that mean z-scores were statistically different between slice-order protocols (VOMOCO: Z = 1.72, p < 0.001) and head-motion correction methods (Interleaved: Z = 1.68, p < 0.001). In the LPC, pairwise comparisons revealed that mean z-scores were statistically different between slice-order acquisitions (SLOMOCO: Z = - 0.9, p = 0.005) and head-motion correction methods (Interleaved: Z = 1.26, p < 0.001).

Considering the whole DMN regions and the PCC node, no statistically significant correlations between absolute mean z-scores and median volume-to-volume displacements were found for any slice-order protocol and head-motion correction

method. Considering bilateral LPC and the ACC regions a weak correlation was found between absolute mean z-scores and median volume-to-volume displacements (LPC/ ACC: $r = 0.3$, $p < 0.05$) only for sequentially acquired data combined with VOMOCO (LPC) and SLOMOCO (ACC).

To summarize, the main characteristic DMN nodes were found and had similar spatial locations in all slice acquisition and head-motion correction methods. However, the group connectivity estimations of the whole DMN and its separate main nodes were significantly affected by slice-order protocol and head-motion correction methods.

4.3.3. TRT reliability metrics

TRT reliability of intrinsic DMN connectivity is reported for each slice-order protocol and head-motion correction method in **Figure 4.5** and summarized in **Table 4.3**. The overall mean (SD) absolute percent errors across subjects, conditions and ROIs were below 8% (DMN: 6.7 ± 0.5%; PCC: 8.0 ± 0.9%; ACC: 6.3 ± 0.4%; LPC: 7.7 ± 0.1%). The overall ICC score averaged across conditions and all ROIs, indicate smoderate mean z-scores reliability (ICC = 0.46; C.I. 0.29 - 0.57); TRT reliability was moderate in the entire DMN (ICC = 0.47; C.I. 0.31 - 0.57), as well as its nodes, PCC (ICC = 0.41; C.I. 0.26 - 0.52), ACC (ICC = 0.49; C.I. 0.32 - 0.62), LPC (ICC = 0.47; C.I. 0.28 - 0.58). There were no statistically significant effects of slice acquisition or motion correction method in the test-retest reliability of mean ROI z-scores.

The Jaccard index indicated lower reliability in the voxel-wise spatial overlap of each functional ROIs with an average of 40 ± 13% (DMN: 39 ± 9%; PCC: 51 ± 15%; ACC: 30 ± 12%; LPC: 39 ± 13%). The Friedman test performed across all subjects, sessions and conditions-of-interest revealed no statistically significant effects of either slice-order acquisition or head-motion correction method in the entire DMN but revealed statistically significant effects in the single nodes of the default-mode system (**Table 4.4**).

When considering the PCC, multiple comparisons applied via pairwise comparisons revealed that mean (SD) TRT reliability were statistically different ($Z = -1.2$, $p = 0.008$) between interleaved (53 ± 8)% and sequential slice-order protocols (62 ± 15)% in SLOMOCO corrected data and between VOMOCO (42 ± 28)% and

SLOMOCO (62 ± 15)% in sequentially acquired data (Z = -1.8, p < 0.001).

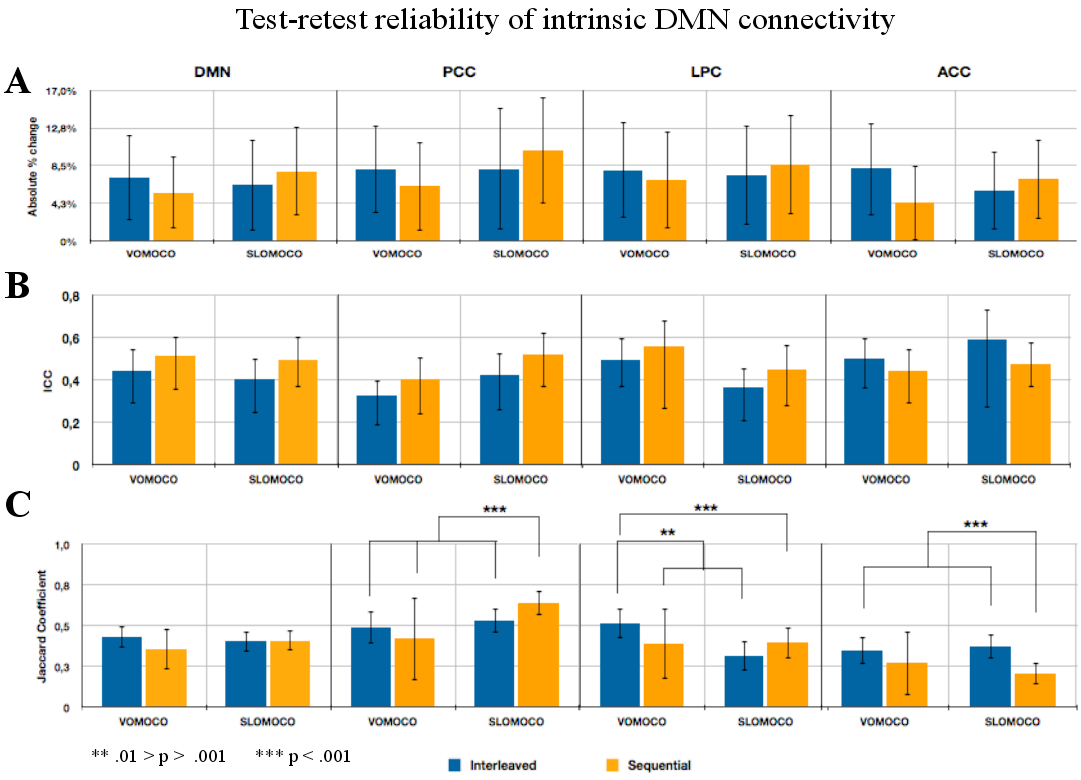

Figure 4.5. Test-retest reliability of intinsic DMN connectivity. The figure shows **(A)** mean (SD) absolute percent error in mean DMN z-scores across subjects; (**B**) intraclass correlation coefficients (C.I.) for mean DMN z-scores; (**C**) mean (SD) Jaccard indexes across subjects for the degree of spatial reproducibility of the DMN. These are shown for each slice-order acquisition sequence, interleaved (blue) and sequential (orange), and motion correction method, VOMOCO and SLOMOCO, for each DMN node under investigation (from left to right: the entire DMN, posterior cingulate cortex (PCC), lateroparietal cortex (LPC), and anterior cingulate (ACC). No statistically significant effects associated with slice-order acquisition or motion correction method were found in the reliability of mean z-scores (moderate) and cluster-overlap (poor) in the DMN. The cluster-overlap was particularly diminished in the ACC. Mixed effects (corrected) were found for the Jaccard index in each single DMN node indicating inter-regional variability of motion correction method and slice-order acquisition effects.

When considering the ACC, mean (SD) TRT reliability was statistically different (Z = 1.9, p < 0.001) between interleaved (37 ± 7)% and sequential slice-order protocols (0.2 ± 0.08) in SLOMOCO corrected data. When considering LPC, mean (SD) TRT reliability was statistically different (Z = 1.17, p < 0.001) between interleaved (52 ± 10)% and sequential (37 ± 22)% slice-order protocols in VOMOCO corrected data and between VOMOCO (52 ± 10)% and SLOMOCO (32 ± 9)% in interleaved acquired data (Z = 1.6, p < 0.001). Importantly, no correlation was found between any TRT reliability

measure and movement-derived metrics, averaged across sessions, in any ROI and for any combination-of-interest.

	Absolute Percent Change				Statistics
	VOMOCO		SLOMOCO		
	Interleaved	Sequential	Interleaved	Sequential	
DMN	7.2 ± 4.8%	5.5 ± 4.1%	6.3 ± 5.2%	7.8 ± 5%	Friedman, $\chi^2(3, n = 23) = 2.5$, $p = .5$
PCC	7.3 ± 5.0%	6.2 ± 5.0%	8.1 ± 6.9%	10.3 ± 6.1%	Friedman, $\chi^2(3, n = 23) = 5.8$, $p = .1$
ACC	8.2 ± 5.2%	4.3 ± 4.2%	5.7 ± 4.4%	7.0 ± 4.5%	Friedman, $\chi^2(3, n = 23) = 56.0$, $p = .06$
LPC	8.0 ± 5.4%	6.9 ± 5.5%	7.4 ± 5.6%	8.7 ± 5.6%	Friedman, $\chi^2(3, n = 23) = 3.7$, $p = .3$
	ICC (C.I.)				
DMN	0.45 (0.29 - 0.55)	0.51 (0.35 - 0.61)	0.4 (0.24 - 0.51)	0.5 (0.37 - 0.61)	
PCC	0.33 (0.19 - 0.4)	0.4 (0.23 - 0.51)	0.43 (0.26 - 0.53)	0.52 (0.37 - 0.63)	
ACC	0.5 (0.36 - 0.6)	0.45 (0.29 - 0.55)	0.59 (0.27 - 0.73)	0.47 (0.36 - 0.58)	
LPC	0.49 (0.36 - 0.6)	0.56 (0.26 - 0.68)	0.36 (0.20 - 0.46)	0.47 (0.28 - 0.57)	

Table 4.3. Reliability of mean DMN z-scores using test-retest and inter-rater reliability approaches. Test-retest percentage errors (mean ± SD) are reported for each condition (columns) and ROI (rows) with pairwise statistical evaluations (uncorrected) across columns (right column). Inter-rather reliability ICC confidence intervals are reported at bottom for comparison. No statistical difference in mean z-scores reliability was found across conditions, ROIs and reliability measures.

To summarize, neither slice-order protocols nor head-motion correction methods affected the across-session TRT reliability of FC-fMRI in single DMN nodes. The Jaccard index emphasizes that the frontal DMN node is the least reliable, with optimal choice of slice acquisition depending on the head-motion correction method. Most importantly, no combination of slice-order protocol and head-motion correction method was found that would offer significant TRT reliability improvements that were systematically consistent across nodes.

	Jaccard index				Statistics
	VOMOCO		SLOMOCO		
	Interleaved	Sequential	Interleaved	Sequential	
DMN	43 ± 7%	36 ± 13%	40 ± 6%	41 ± 6%	Friedman, $\chi^2(3, n=23)=6.8$, $p = .08$
PCC	49 ± 10%	44 ± 30%	53 ± 8%	64 ± 8%	Friedman, $\chi^2(3, n = 23) = 29.8$, $p < .001$
ACC	35 ± 8%	29 ± 28%	37 ± 7%	21 ± 7%	Friedman, $\chi^2(3, n = 23) = 31.5$, $p < .001$
LPC	52 ± 10%	39 ± 20%	32 ± 9%	40 ± 9%	Friedman, $\chi^2(3, n = 23) = 20.6$, $p < .001$

Table 4.4. Summary of spatial inter-session convergence of the DMN. Jaccard index (mean ± SD) are shown for each condition (column) and ROI (rows). Non-parametric statistical comparisons are also reported in the last column on the right (uncorrected). No significant statistical difference was found in the Jaccard index across conditions in the entire DMN. However, mixed effects were instead detected at the level of each single DMN ROI. The magnitude of spatial reproducibility was the lowest in the ACC.

4.4. Limitations

This present work suggests that intrinisc DMN connectivity estimates derived using model-free methods are sensitive to data acquisition and analysis choices specifically related to head-motion issues. However, the overall head-motion detected was low in this sample of collaborative young healthy adults. This might have hidden important slice-order protocol and/or motion correction method effects in the TRT reliability of the DMN. Moreover, this also prevented us from comparing our factors as a function of the amount of head-motion in the data. It would be interesting to investigate these issues in old individuals and children who are intrinsically characterized by higher motion and reduced connectivity in the anterior node of the DMN.

Of note, one limitation of this study is that slice-order protocols were not counterbalanced across MRI scanning sessions with interleaved always preceeding sequential acquisitions. Slice-order acquisition counterbalancing may be critical to eliminate biases induced by fatigue effects in the BOLD signal or intrinsic DMN connectivity measures. Even though, tSNR examinations did not reveal any significant difference associated with slice-order acquisition sequences in brain volumes prior

motion correction, slice-order counterbalancing across MRI sessions would optimize the overall study design in future longitudinal studies.

This study does not provide information about the TRT reliability of model-based methods such as graph-theoretical approaches. These have been demonstrated to have comparable reliability with model-free intrinsic FC-fMRI methods. The present scenario has shown overall low head-motion contamination which favors comparisons across these two methods not only regarding the DMN but also to other intrinsic brain networks with potential evaluations of between-network interactions.

5

Discussion

5.1. Main findings

After having outlined the two main studies of this thesis, their results will be compared in this section and will be separately discussed and put in perspective with the literature in the following sections.

Despite differences in subject populations and MRI equipment used, the results of our studies indidate that the intrinsic DMN connectivity is significantly sensitive to choices in slice acquisition, physiological denoising and head motion correction methods but that the DMN reliability is not, or at least not systematically. This is probably due to technical properties of the ICA methdology to consistently detect the DMN despite our manipulations. This encourages the usage of group-ICA and dual-regression in longitudinal intrinsic FC-fMRI studies under a variety of acquisition and analysis choices.

Table 5.1. compares acquisition and analysis protocols between the two studies and relative DMN reliability results. They markedly differ in acquisition protocol choices but both underwent similar preprocessing workflows. For ease of comparison, only the shared preprocessing workflow that includes VOMOCO and FIX-standard physiological denoising was considered. The order of preprocessing steps was identical between studies. One main difference is in the application of spatial smoothing, which was avoided in the head motion study. Both studies used the group-ICA and dual-regression methodology followed by identical cluster analyses to characterize the DMN.

However, the two studies resampled voxels at different spatial resolution during MNI coregistration and conducted group-ICA at different dimensionalities. Group-ICA dimensionality is characteristic of each single dataset. Our empirical evaluations indicated that low dimensionality is associated with higher DMN reliability in datasets characterized by high variability and small sample size (N=10 per site) (Jovicich et al.,

2016; Marchitelli et al., 2016). In contrast, single-site studies with larger sample size (N=47) can benefit from a higher number of ICs for decomposition.

	Physiological denoising study	***Head-motion correction study***
Data acquisition		
Field Strength Manufacturer	3T multisite (13) (Siemens, Philips, GE)	4T single site Bruker Medspec
Participants	N = 65 (62.6 ± 5.3y)	N = 24 (27 ± 5.3y)
TRT interval	17 ± 10.2d	4.2 ± 4.4d
TE/TR	30 ms / 2.7s	30 ms / 2.2s
Slice acquisition (axial)	40 Interleaved 0.45 mm slice-gap	37 Interleaved / sequential 0.6 mm slice-gap
Preprocessing		
Motion correction	3D	3D
Physiological denoising method	FSL-FIX (standard)	FSL-FIX (standard)
Spatial Smoothing	6 mm FWHM	N.A.
Group analysis		
Spatial MNI resolution	4x4x4 mm	3x3x3 mm
ICA dimensionality	10 IC	15 IC
DMN extraction	Group-ICA / dual-regression	Group-ICA / dual-regression
TRT reliability of DMN connectivity		
Absolute % change	(9.7 ± 9)%	(6.3 ± 4.5)% *
Jaccard index	(60 ± 15)%	(40 ± 10)% ***

Table 5.1. Comparison between data acquisition, analysis choices and DMN reliability results across the two studies. The table compares the two studies (second, third columns) for acquisition, analysis (preprocessing, postprocessing) and TRT reliability results (first column). For ease of comparison, the two studies are considered only for preprocessing workflows that follow volume-based head-motion correction and FIX-standard physiological denoising. Main differences include field strength, number of sites, age, time interval across scanning repetitions, TR, number of slices and slice acquisition methods, spatial resolution after spatial normalization, number of ICs. N.A. = not applied. * = $p < 0.05$, *** = $p < 0.001$;

The absolute percent change of mean z-scores averaged across all subjects from the 3T multisite study (N=65) and across all subjects and slice acquisition methods from

the 4T single-site study (N=47) were both below 10%. However, an independent samples t-test indicates higher reliability in the young cohort (6.3 ± 4.5)% when compared with old individuals (9.7 ± 9)%; $t(110) = -2.16$, $p < .05$. No regional variation was found in the reliability scores among the DMN nodes in the young cohort. In contrast, previous studies showed increased regional variability in healthy elders (Jovicich et al., 2016). The large variability between these two studies for MRI hardware and study-design makes it challenging to attribute this difference merely to demographic or cognitive factors. However, we feel confident to speculate that the strength of DMN connectivity is more stable over time in younger individuals at shorter intervals between MRI scan repetitions.

Despite the high TRT reliability measured in the young cohort, the corresponding ICC indicates only moderate / fair reliability (mean ICC: 0.47). This finding agrees with previous DMN reliability studies (Shehzad et al., 2009; Braun et al., 2012; Guo et al., 2012) who reported comparable ICCs (range: 0.2, 0.52) and is driven by similar within- and between-subject variances in mean z-scores. This might indicate that ICC analyses show reduced sensitivity to aging effects. Indeed, previous studies indicate comparable ICCs between young and old individuals using either model-based or model-free connectivity methods (Song et al., 2012; Guo et al., 2012). Furthermore, it was also shown that ICC analyses might be sensitive to sample size (Bennet & Miller, 2010) and FC-fMRI method choices, with group-ICA and dual-regression approaches yielding the highest reliability in healthy aging (Jovicich et al., 2016; Guo et al., 2012). In the physiological noise correction study, due to the small sample size at each site, the ICC analysis for the DMN connectivity measurements could not be performed and was conducted only to examine the inter-site “agreement” in the reliability scores.

When considering the DMN spatial reproducibility within subjects, an independent-samples t-test indicates statistically significant higher spatial overlap in old individuals (60 ± 15)% than young adults (39.5 ± 9.8)%; $t(110) = 8.1$, $p < .001$. There was regional variability in Jaccard indexes within the DMN, with the PCC (52 ± 13)% and the ACC (30 ± 11)% being the most and the least reliable nodes, respectively. Higher spatial overlap in the PCC was also found in healthy old individuals (Jovicich et al., 2016) while other works showed that the ACC is generally characterized by higher spatial variability (Van Dijk et al., 2010).

One possibility of this difference across the two studies is that DMN signals become less detectable as field strenght increases or that spatial smoothing is essential to improve signal detection at higher field strength. Differences in ICA settings could have also led to these differences, with lower Jaccard indexes associated with higher IC dimensionality and higher spatial resolution after spatial MNI normalization. Future investigations should address these issues.

Finally, the magnitude of spatial reproducibility indicates that the Jaccard index would not be appropriate to validate spatial DMN biomarkers. It should be noted that these reliability measures are unitless and other affine metrics, such as the cosine similarity coefficient (Thada & Jaglan, 2013), could be more adequate to this purpose, yielding higher reliability in the old cohort (75 ± 12)%.

5.2. DMN connectivity in aging: MRI-site and physiological denoising effects

In the first study, the influence of some rPNC methods was investigated on the TRT reliability of intrinsic DMN connectivity measured in elderly participants across a consortium of 13 MRI sites. As a recent study showed (Jovicich et al., 2016), despite MRI acquisition harmonization efforts, this consortium showed high variability of tSNR across sites. In this study those findings were extended to show that, in spite of the high tSNR variability, the site-group DMN was able to be automatically selected for all rPNC methods, allowing robust DMN characterization at the single-subject level. This suggests that the rPNC methods evaluated here are valid candidates for longitudinal multisite RS-fMRI studies with elderly participants. However, the overlap measure using a full DMN-template was not always successful at automatically detecting the DMN among the components at every site and rPNC condition. The accuracy of the automatic selection procedure improved only once frontal ROIs were excluded from the template. This was due to high variability of frontal DMN areas for co-activation cluster-size (number of active voxels), spatial coordinate locations of active voxels in relation to the template, and the presence of large structured noise in these regions in other components. The FC-fMRI variability in frontal regions might be a sign of age-related decline of mean z-scores in the DMN of healthy aging subjects (Damoiseaux et al., 2008; Huang et al., 2015).

An intrinsic challenge of characterizing resting-state networks with ICA methods is the possibility of network splitting (Abou Elseoud et al., 2011), which can also affect reliability (Zuo & Xing, 2014). In this study, the DMN splitting phenomenon was addressed in several ways: first, the number of ICs were kept low and constant for all sties (Abou Elseoud et al., 2011). Second, a supervised algorithm was used to select the group component with maximum overlap to a DMN reference template and the group DMN results were visually inspected. As with most studies, it is possible that some components of the DMN were not properly included in our estimations, and instead were distributed through one or more of the other rejected components. However, since the rPNC methods were applied at the single-subject level before group-ICA, after which all rPNC methods followed a similar workflow, the comparison of rPNC methods was not expected to be strongly biased by DMN splitting effects.

It was found that, among the rPNC methods, FSL-FIX increased mean z-scores in the DMN without affecting the DMN cluster-size in many sites of the consortium. This finding is potentially due to FSL-FIX, which removes a larger proportion of signal-of-no-interest from the data than the other rPNC methods (such as residual motion artifacts), leading to an increase in mean z-scores within the DMN. This effect was found also in other resting-state networks (Marchitelli et al., 2016). This suggests that, the application of FSL-FIX could increase the sensitivity of group-ICA to detect the DMN in healthy aging subjects and discern abnormal connectivity patterns in clinical aging, a task known to be not straightforward (Bai et al., 2008; Koch et al., 2012). Importantly, training the classifier at each site improved MRI site consistency in DMN-derived measurements and confined individual spatial DMN maps within converging spatial locations across MRI sites. Relative to NPC, FSL-FIX was found to increase the number of active voxels in gray matter areas outside the DMN nodes, together with surrounding white matter and CSF voxels (Marchitelli et al., 2016). This finding stresses the importance of the selection of nodes that will be defined as members of the DMN.

These findings are consistent with previous works showing that DMN connectivity estimates strongly depend on how thoroughly physiological and other confounds are removed upstream (Beall & Lowe, 2010; Birn, 2012; Murphy et al., 2013). Of note, in this work only ICA (with a specific number of components) was

considered to characterize the DMN. This choice was mainly motivated by the high DMN reliability obtained via ICA in the previous work (Jovicich et al., 2016).

5.3. DMN reliability in aging: MRI-site and physiological denoising effects

The TRT reliability of DMN connectivity metrics was consistent across all MRI sites for all rPNC methods, which allowed the averaging across sites to estimate consortium reliability metrics. Absolute percent errors were in the range of 5-11% for DMN z-score and cluster-size reliability. Jaccard indexes were in the range 60-65% for DMN cluster-overlap reliability. Only for FSL-FIX an inverse correlation between motion and Jaccard indexes was found. This dependence might originate from the additional cleaning of residual motion-related spatial maps implemented by FSL-FIX, which leads to higher characterization of subject-specific DMN maps when the overall motion detected is lower.

These results are consistent with similar measures recently reported in a single-site study with young healthy participants (Meindl et al., 2010). In the present study with elderly participants, high reliability was found for DMN cluster-size and moderate reliability for DMN cluster-overlap, consistent across sites, albeit not influenced by rPNC methods. The statistical evaluations of MRI site effects in TRT reliability scores were performed using both parametric and non-parametric approaches: both tests agreed in the lack of MRI site effects for DMN mean z-scores and cluster-size reliability. DMN cluster-overlap reliability results differed across tests, giving no MRI site effects with the non-parametric test but significant MRI site effects with the 2-way ANOVA. Given the low number of subjects it is difficult to interpret this difference since the non-parametric test is robust to outlier but underpowered with low samples while the 2-way ANOVA has the opposite characteristics.

In particular, no rPNC correction method showed significant DMN reliability improvements relative to NPC. These findings agree with those from a recent single-site study (Birn et al., 2014) that found that neither correction based on parallel measures (cardiac and respiratory measures) nor Tissue-based correction improved the TRT reliability of mean z-scores among ROIs, including the DMN in young adults. This

study extends those results in several ways: population age, multicentric MRI consortium and rPNC methods.

To the best of our knowledge, this is the first study that investigates the influence of some rPNC methods on the consistency of DMN connectivity reliability in a multisite study. In fact, although DMN reliability is consistent across sites, rPNC methods are still likely to influence the degree of reproducibility consistency across the consortium.

When considering mean z-scores in the DMN, Tissue-based and FIX-standard correction methods significantly improved the reproducibility consistency of the DMN amongst the different 3T MRI sites. This effect may be driven by exploiting individual anatomical information (T1-anatomical) to remove physiological noise from non-gray matter tissue. This suggests that the use of single-site derived spatial priors in PESTICA may not be sufficient to contrast multisite heterogeneity. When considering reproducibility consistency, FIX-training did not perform as well as FIX-standard. This may be related to the deviation of non-brain related percent variance classified between the two methods (FIX-training retained higher signal variance, very likely not to be associated to the DMN). Furthermore, such variability was unequally distributed across sites, leading to overall lower reproducibility consistency of the DMN relative to FIX-standard. Therefore, we recommend adopting a higher signal-noise threshold for FSL-FIX in longitudinal multisite resting-state FC-fMRI studies.

5.4. DMN connectivity: slice-order acquisition and motion correction effects

The main goal of the second reliability study was to evaluate whether different combinations of slice-order fMRI acquisition protocols and head-motion correction techniques would differently affect the TRT reliability of intrinsic DMN connectivity. Subjects were instructed to stay still to investigate these effects in standard experimental conditions rather than with artificially directed movement. Using group-ICA and dual-regression methods the group DMN and reconstructed individual DMN components were all identified under all slice-order protocols and head-motion correction methods in both fMRI sessions. These findings highlight the reliable nature of the ICA approach

(Beckmann et al., 2009) to define consistent DMN maps across subjects and sessions (Chen et al., 2008; Damoiseaux et al., 2006).

We found that both slice-order protocols and head-motion correction methods significantly affect mean z-scores within the DMN. These effects are not likely due to fatigue effects considering that no differences in the tSNR were found across slice-order protocols before head-motion correction. It is more likely that this might depend on several factors: first, on the variability between head-motion correction methods to detect and remove the motion component of the BOLD signal. In fact, while head-motion estimated using VOMOCO indicated overall low movement, significant more movement was estimated after SLOMOCO. In average across all participants and sessions (N=47), SLOMOCO metrics registered 64% and 61% higher motion than VOMOCO under interleaved and sequential slice-order protocols, respectively. This finding is in line with (Beall & Lowe, 2014) who showed that VOMOCO estimates miss on average at least the 50% of non-volumetric motion in real data. Second, on BOLD effects arising from blood inflow effects, i.e. magnetically saturated blood flows from slice to slice, that depend on the ascending sequential slice-order protocol adopted. Although an earlier task-based 2T fMRI study demonstrated that no BOLD blood saturation issue arises from the selection of different slice-order protocols and that activations in the visual cortex overlap quite well across different slice-order protocols (Turner et al., 1998), these results suggest that slice-order selection cannot be an underestimated issue in RS-fMRI connectivity studies of the DMN at 4T.

Both slice-order protocol and head-motion correction method effects were found within the entire DMN as well as its main nodes but were less pronounced in the PCC region. In fact, in contrast to all other DMN nodes, PCC connectivity was very robust against potential surviving motion artifacts showing no correlations with the estimated motion and no head-motion correction method effects. These observations confirm the resilience of the main hub of the DMN (Fransson & Marrelec, 2008) and support research in PCC-based connectivity markers of dementia (Bai et al., 2009; Zhou et al., 2008).

5.5. DMN reliability: slice-order acquisition and motion correction effects

TRT reliability measures of mean DMN z-scores in our group of healthy young participants indicate overall moderate-to-good reliability with absolute percent error below 8% and ICC range of 0.29 - 0.57. These results are in line with previous DMN reliability studies who adopted interleaved slice acquisition and VOMOCO methods, implementing the same methodology (i.e. Group-ICA and dual regression) in healthy young adults (Zuo et al., 2010) and healthy old individuals (Guo et al., 2012).

In a previous study, it was shown that there is dependence between the amount of head-motion and DMN reliability in healthy old individuals, with high motion associated with reduced ICC (Guo et al., 2012). In the present study both reliability measures did not correlate with motion-derived metrics and were not influenced by both slice-order acquisition and motion correction methods in the short-term. This might primarily depend on our sample of young healthy adults who are characterized by low head-motion and were instructed not to move during scanning sessions. Investigating the influence of these combinations on the TRT reliability of FC-fMRI in high-motion populations will deepen our knowledge about potential optimizations in acquisition sequences and data analysis choices.

Also, model-free ICA-based approaches are known to characterize highly reproducible components (Zuo et al., 2010). The adoption of this approach might therefore explain the present outcome even though a comparison with model-based approaches (i.e. seed-based methods) was not directly investigated in the present study. Of note, model-free ICA-based methods were here optimized to circumvent their unsupervised nature (Abou Elseoud et al., 2011; Calhoun et al., 2009) and ICA implementations were twofold: first, during individual image preprocessing using FSL-FIX (Griffanti et al., 2014; Salimi-Khorshidi et al., 2014) to remove additional physiological motion and hardware-related artifacts and second, during group post-processing using group-ICA to characterize the DMN and isolate its component from spatio-temporal patterns of non-neural origins.

The voxel-wise spatial overlap of within-subject DMN maps was lower than mean z-score reliabilities, indicating an overall spatial DMN reproducibility of 40% in agreement with a recent DMN reliability study of healthy young adults (Meindl et al.,

2010) who reported Dice coefficients for spatial DMN reproducibility of 42% (J = 33%) at 1x1x1 spatial voxel resolution using 3T MRI scanner. Spatial reproducibility was the highest in the PCC (51%) and the lowest in the ACC (30%), consistently with previous reported findings (Jovicich et al., 2016; Van Dijk et al., 2010). Neither slice acquisition nor head-motion correction methods influenced the voxel-wise spatial reproducibility of the entire DMN.

In contrast, Jaccard indexes were sensitive to these factors in the single nodes of the DMN. In particular none of the slice acquisition and motion correction methods improved reliability in the ACC which was slightly lower than in (Meindl et al., 2010), ACC (J = 47%). Higher spatial variability in the ACC was already reported in other 1.5T and 3T studies (Damoiseaux et al., 2006; Meindl et al., 2010) across subjects and sessions. Altogether, these results have important implications for the investigations of ACC-based longitudinal connectivity markers of psychopathological conditions (Davey et al., 2012; Tian et al., 2006).

Future investigations of descending slice-order protocols might help clarify whether this instability in the ACC activation cluster-size might be minimized using sequential acquisitions, thought of being more robust against blood inflow saturation effects in that direction with overall low sensitivity to motion spin-history artifacts (http://imaging.mrc-cbu.cam.ac.uk/imaging/TipsForDataAcquisition#Slice_order). Most importantly, no combination of slice-order protocols and head-motion correction methods improved the TRT reliability of FC-fMRI consistently across all DMN nodes. The present study therefore suggests relative freedom of choice between the examined slice-order protocols and head-motion correction methods when using model-free methods in longitudinal studies of intrinisc DMN connectivity.

5.6. Future studies

The analyses presented in this thesis can be extended on several fronts to further investigate the TRT reliability of intrinsic DMN connectivity. The head-motion study showed that different reliability metrics might not perfectly converge. This indicates that some metrics can bias the interpretation of reliability and limit the comparability and generalizability of findings. In particular, the ICC is limited by its specific

dependence on the data set under investigation (Bennet & Miller, 2010). In the head-motion study, despite the low absolute percent change in DMN mean z-scores, low between-subject variability could have biased the ICC analysis. Importantly, the conspicuous variability in ICC models adopted in the literature makes it challenging to interpret the reliability of the functional connectome. Many previous studies adopted ICC models ascribable to the following:

$$\mathrm{ICC}(1,1) = \frac{\mathrm{BMS} - \mathrm{WMS}}{\mathrm{BMS} + (k-1)\mathrm{WMS}}$$

with BMS indicating the mean square between subjects and WMS the mean square within subject, and k the number of fMRI sessions. This ICC model admits negative ICC values that cannot be easily interpreted (Müller & Büttner, 1994; Rousson et al., 2002) and are tipically set to zero (Kong et al., 2007; Zhang et al., 2011). In agreement with (Bennett & Miller, 2010), we simplified the ICC equation by dividing the variance of interest by the total variance:

$$\mathrm{ICC}(1,1) = \frac{\mathrm{BMS}}{\mathrm{BMS} + (k-1)\mathrm{WMS}}$$

in this equation, a value of 0 would only indicate that there is no agreement between the values across the k sessions, *because within-subject variability would dominate the equation* (Bennett & Miller, 2010). Direct comparisons of these ICC formulas and the evaluation of multiple TRT reliability measures are still needed in the context of unweighted and undirected graphs (i.e. standard measures of FC-fMRI). This is necessary to improve the current reliability framework in the context of effective connectivity measures of the human brain that yield directed graphs to explore causal relationships between several brain systems (Li et al., 2009).

To this aim, ICA-based functional activation clusters defined in this thesis could be used as spatial priors or reference for directed graph-based connectivity approaches to define the DMN (Franco et al., 2013). The longitudinal stability of the spatial DMN patterns demonstrates that ICA methods are valid and clear methodological criteria to define apriori DMN regions and are thus suitable to both serve as control and investigate abnormal resting-state FC-fMRI in patients (Fox & Greicius, 2010). Considering the wide set of graph-based connectivity properties and the relative degree

of variability in TRT reliability methods, group-ICA DMN reliability constitutes an optimal term of comparison for longitudinal evaluations of directed graph-based connectivity approaches. Moreover, in the head-motion study, DMN reliability was comparable between interleaved and sequential slice-order acquisition methods. The influence that these acquisition choices might exert on effective connectivity methods such as dynamic causal modeling (DCM) were postulated but not addressed in a systematic and empiric way (Stephan et al., 2010).

Importantly, intrinsic brain networks besides the DMN might also show clinical relevance and deserve reliability investigations. The salience network is involved in fronto-temporal dementia (Seeley et al., 2007, 2008) and chronic pain (Greicius et al., 2008; Cauda et al., 2009) and interacts with the amygdala and the frontoparietal control network in generalized anxiety disorder (Etkin et al., 2009) and with the DMN in ADHD (Weng et al., 2010). Therefore future reliability studies should not only consider measures of intrinsic connectivity within some brain networks but also between different brain systems (Wang et al., 2015).

From our experience in longitudinal multisite resting-state FC-fMRI in healthy aging, it turns out that stringent correction for multiple comparisons will be very important in clinical studies and that correlations with clinical variables are needed to account for between-subject variability. Our investigations into the impact of different physiological denoising approaches in healthy old individuals should be translated to patients, together with the investigations of other preprocessing choices (Fox & Greicius, 2010). Although head-motion does not influence ICA-based connectivity and its reliability, it is always necessary to analyze motion metrics in healthy controls and patients. We believe that reliability in some clinical populations might be sensitive to motion and the stability of FC-fMRI measures might improve after 2D slice-based motion correction methods also in patients.

Novel methods that make use of reliability indexes to further correct FC-fMRI in intrinsic brain networks are becoming available (Mueller et al., 2015). These reliability-based correction methods will improve the comparability between healthy controls and patients.

6

References

Abou Elseoud, A., Littow, H., Remes, J., Starck, T., Nikkinen, J., Nissilä, J., Timonen, M., Tervonen, O., Kiviniemi, V. (2011). Group-ICA Model Order Highlights Patterns of Functional Brain Connectivity. *Front Syst Neurosci,* 5:37.

Abou-Elseoud, A., Starck, T., Remes, J., Nikkinen, J., Tervonen, O., & Kiviniemi, V. (2010). The effect of model order selection in group PICA. *Hum Brain Mapp, 31*(8), 1207-1216.

Andreasen, N.C., O'Leary, D.S., Cizadlo, T., Arndt, S., Rezai, K., Watkins, G.L., Ponto, L.L., Hichwa, R.D. (1995). Remembering the past: two facets of episodic memory explored withpositron emission tomography. Am J Psychiatry. 152(11), 1576-85.

Andrews-Hanna, J. R. (2012). The brain's default network and its adaptive role in internal mentation. *Neuroscientist, 18*(3), 251-270.

Andrews-Hanna, J. R., Snyder, A. Z., Vincent, J. L., Lustig, C., Head, D., Raichle, M. E., & Buckner, R. L. (2007). Disruption of large-scale brain systems in advanced aging. *Neuron, 56*(5), 924-935.

Andronache, A., Rosazza, C., Sattin, D., Leonardi, M., D'Incerti, L., Minati, L., & Coma Research Centre - Besta, I. (2013). Impact of functional MRI data preprocessing pipeline on default-mode network detectability in patients with disorders of consciousness. *Front Neuroinform, 7*, 16.

Anticevic, A., Cole, M. W., Murray, J. D., Corlett, P. R., Wang, X. J., & Krystal, J. H. (2012). The role of default network deactivation in cognition and disease. *Trends Cogn Sci, 16*(12), 584-592.

Arthurs, O. J., & Boniface, S. (2002). How well do we understand the neural origins of the fMRI BOLD signal? *Trends Neurosci, 25*(1), 27-31.

Bai, F., Watson, D. R., Yu, H., Shi, Y., Yuan, Y., & Zhang, Z. (2009). Abnormal resting-state functional connectivity of posterior cingulate cortex in amnestic type mild cognitive impairment. *Brain Res, 1302*, 167-174.

References

Bai, F., Zhang, Z., Yu, H., Shi, Y., Yuan, Y., Zhu, W., Zhang, X., Qian, Y. (2008) Default-mode network activity distinguishes amnestic type mild cognitive impairment from healthy aging: a combined structural and resting-state functional MRI study. Neurosci Lett, 438(1), 111-5.

Balthazar, M. L., de Campos, B. M., Franco, A. R., Damasceno, B. P., & Cendes, F. (2014). Whole cortical and default mode network mean functional connectivity as potential biomarkers for mild Alzheimer's disease. *Psychiatry Res, 221*(1), 37-42.

Beall, E. B. (2010). Adaptive cyclic physiologic noise modeling and correction in functional MRI. *J Neurosci Methods, 187*(2), 216-228.

Beall, E. B., & Lowe, M. J. (2007). Isolating physiologic noise sources with independently determined spatial measures. *Neuroimage, 37*(4), 1286-1300.

Beall, E. B., & Lowe, M. J. (2010). The non-separability of physiologic noise in functional connectivity MRI with spatial ICA at 3T. *J Neurosci Methods, 191*(2), 263-276.

Beall, E. B., & Lowe, M. J. (2014). SimPACE: generating simulated motion corrupted BOLD data with synthetic-navigated acquisition for the development and evaluation of SLOMOCO: a new, highly effective slicewise motion correction. *Neuroimage, 101*, 21-34.

Beckmann, C. F., DeLuca, M., Devlin, J. T., & Smith, S. M. (2005). Investigations into resting-state connectivity using independent component analysis. *Philos Trans R Soc Lond B Biol Sci, 360*(1457), 1001-1013.

Beckmann, C. F., & Smith, S. M. (2004). Probabilistic independent component analysis for functional magnetic resonance imaging. *IEEE Trans Med Imaging, 23*(2), 137-152.

Beckmann, M., Filippini, & Smith. (2009). Group comparison of resting-state FMRI data using multi-subject ICA and dual regression. *Neuroimage, 47*(S148).

Behzadi, Y., Restom, K., Liau, J., & Liu, T. T. (2007). A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. *Neuroimage, 37*(1), 90-101.

Bell, A. J., & Sejnowski, T. J. (1995). An information-maximization approach to blind separation and blind deconvolution. *Neural Comput, 7*(6), 1129-1159.

Bennett, C.M., & Miller, M.B. (2010). How reliable are the results from functional magnetic resonance imaging? Ann NY Acad Science, 1191,133-55.

Bianciardi, M., Fukunaga, M., van Gelderen, P., Horovitz, S. G., de Zwart, J. A., Shmueli, K., & Duyn, J. H. (2009). Sources of functional magnetic resonance imaging signal fluctuations in the human brain at rest: a 7 T study. *Magn Reson Imaging, 27*(8), 1019-1029.

References

Binder, J. R., Frost, J. A., Hammeke, T. A., Bellgowan, P. S., Rao, S. M., & Cox, R. W. (1999). Conceptual processing during the conscious resting state. A functional MRI study. *J Cogn Neurosci, 11*(1), 80-95.

Binnewijzend, M.A., Schoonheim, M.M., Sanz-Arigita, E., Wink, A.M., van der Flier, W.M., Tolboom, N., Adriaanse, S.M., Damoiseaux, J.S., Scheltens, P., van Berckel, B.N., Barkhof, F. (2012) Resting-state fMRI changes in Alzheimer's disease and mild cognitive impairment. *Neurobiol Aging*, 33(9), 2018-28.

Biomarkers Definitions Working, G. (2001). Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. *Clin Pharmacol Ther, 69*(3), 89-95.

Birn, R. M. (2012). The role of physiological noise in resting-state functional connectivity. *Neuroimage, 62*(2), 864-870.

Birn, R.M., Cornejo, M.D., Molloy, E.K., Patriat, R., Meier, T.B., Kirk, G.R., Nair, V.A., Meyerand, M.E., Prabhakaran, V. (2014). The influence of physiological noise correction on test-retest reliability of resting-state functional connectivity. *Brain Connect,* 4(7), 511-22.

Birn, R. M., Diamond, J. B., Smith, M. A., & Bandettini, P. A. (2006). Separating respiratory-variation-related fluctuations from neuronal-activity-related fluctuations in fMRI. *Neuroimage, 31*(4), 1536-1548.

Birn, R. M., Murphy, K., & Bandettini, P. A. (2008). The effect of respiration variations on independent component analysis results of resting state functional connectivity. *Hum Brain Mapp, 29*(7), 740-750.

Birn, R. M., Murphy, K., Handwerker, D. A., & Bandettini, P. A. (2009). fMRI in the presence of task-correlated breathing variations. *Neuroimage, 47*(3), 1092-1104.

Biswal, B., Yetkin, F. Z., Haughton, V. M., & Hyde, J. S. (1995). Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. *Magn Reson Med, 34*(4), 537-541.

Biswal, B. B., & Kannurpatti, S. S. (2009). Resting-state functional connectivity in animal models: modulations by exsanguination. *Methods Mol Biol, 489*, 255-274.

Biswal, B.B., Mennes, M., Zuo, X.N., Gohel, S., Kelly, C., Smith, S.M., Beckmann, C.F., Adelstein, J.S., Buckner, R.L., Colcombe, S., Dogonowski, A.M., Ernst, M., Fair, D., Hampson, M., Hoptman, M.J., Hyde, J.S., Kiviniemi, V.J., Kötter, R., Li, S.J., Lin, C.P., Lowe, M.J., Mackay, C., Madden, D.J., Madsen, K.H., Margulies, D.S., Mayberg, H.S., McMahon, K., Monk, C.S., Mostofsky, S.H., Nagel, B.J., Pekar, J.J., Peltier, S.J., Petersen, S.E., Riedl, V.,

Rombouts, S.A., Rypma, B., Schlaggar, B.L., Schmidt, S., Seidler, R.D., Siegle, G.J., Sorg, C., Teng, G.J., Veijola, J., Villringer, A., Walter, M., Wang, L., Weng, X.C., Whitfield-Gabrieli, S., Williamson, P., Windischberger, C., Zang, Y.F., Zhang, H.Y., Castellanos, F.X., Milham, M.P. (2010). Toward discovery science of human brain function. *Proc Natl Acad Sci U S A*, 107(10), 4734-9.

Biswal, B. B., Van Kylen, J., & Hyde, J. S. (1997). Simultaneous assessment of flow and BOLD signals in resting-state functional connectivity maps. *NMR Biomed, 10*(4-5), 165-170.

Blautzik, J., Vetter, C., Peres, I., Gutyrchik, E., Keeser, D., Berman, A., Kirsch, V., Mueller, S., Pöppel, E., Reiser, M., Roenneberg, T., Meindl, T. (2013). Classifying fMRI-derived resting-state connectivity patterns according to their daily rhythmicity. *Neuroimage*, 71, 298-306.

Blautzik, J., Vetter, C., Schneider, A., Gutyrchik, E., Reinisch, V., Keeser, D., Paolini, M., Pöppel, E., Bao, Y., Reiser, M., Roenneberg, T., Meindl, T. (2014). Dysregulated daily rhythmicity of neuronal resting-state networks in MCI patients. *Chronobiol Int*, 31(9), 1041-50.

Boubela, R. N., Kalcher, K., Huf, W., Kronnerwetter, C., Filzmoser, P., & Moser, E. (2013). Beyond Noise: Using Temporal ICA to Extract Meaningful Information from High-Frequency fMRI Signal Fluctuations during Rest. *Front Hum Neurosci, 7*, 168.

Braun, U., Plichta, M.M., Esslinger, C., Sauer, C., Haddad, L., Grimm, O., Mier, D., Mohnke, S., Heinz, A., Erk, S., Walter, H., Seiferth, N., Kirsch, P., Meyer-Lindenberg, A. (2012). Test-retest reliability of resting-state connectivity network characteristics using fMRI and graph theoretical measures. *Neuroimage*, 59(2), 1404-12.

Brewer, J. A., Worhunsky, P. D., Gray, J. R., Tang, Y. Y., Weber, J., & Kober, H. (2011). Meditation experience is associated with differences in default mode network activity and connectivity. *Proc Natl Acad Sci U S A, 108*(50), 20254-20259.

Bright, M. G., & Murphy, K. (2013). Removing motion and physiological artifacts from intrinsic BOLD fluctuations using short echo data. *Neuroimage, 64*, 526–537.

Buckner, R. L. (2012). The serendipitous discovery of the brain's default network. *Neuroimage, 62*(2), 1137–1145.

Buckner, R. L., Andrews-Hanna, J. R., & Schacter, D. L. (2008). The brain's default network: anatomy, function, and relevance to disease. *Ann N Y Acad Sci, 1124*, 1-38.

Buckner, R. L., & Carroll, D. C. (2007). Self-projection and the brain. *Trends Cogn Sci, 11*(2), 49-57.

References

Buckner, R. L., & Vincent, J. L. (2007). Unrest at rest: default activity and spontaneous network correlations. *Neuroimage, 37*(4), 1091-1096; discussion 1097-1099.

Bullmore, E., & Sporns, O. (2009). Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat Rev Neurosci, 10*(3), 186-198.

Bulmer, M. G. (1979). Principles of statistics. p. 63.

Cauda, F., D'Agata, F., Sacco, K., Duca, S., Cocito, D., Paolasso, I., Isoardo, G., and Geminiani, G. (2009). Altered resting state attentional networks in diabetic neuropathic pain. *Journal of Neurology, Neurosurgery & Psychiatry, jnnp-2009*.

Calhoun, V. D., & Adali, T. (2012). Multisubject independent component analysis of fMRI: a decade of intrinsic networks, default mode, and neurodiagnostic discovery. *IEEE Rev Biomed Eng, 5*, 60-73.

Calhoun, V. D., Adali, T., McGinty, V. B., Pekar, J. J., Watson, T. D., & Pearlson, G. D. (2001a). fMRI activation in a visual-perception task: network of areas detected using the general linear model and independent components analysis. *Neuroimage, 14*(5), 1080-1088.

Calhoun, V. D., Adali, T., Pearlson, G. D., & Pekar, J. J. (2001b). A method for making group inferences from functional MRI data using independent component analysis. *Hum Brain Mapp, 14*(3), 140-151.

Calhoun, V. D., Liu, J., & Adali, T. (2009). A review of group ICA for fMRI data and ICA for joint inference of imaging, genetic, and ERP data. *Neuroimage, 45*(1 Suppl), S163-172.

Chang, C., Cunningham, J. P., & Glover, G. H. (2009). Influence of heart rate on the BOLD signal: the cardiac response function. *Neuroimage, 44*(3), 857-869.

Chen S, Ross TJ, Zhan W, Myers CS, Chuang KS, Heishman SJ, Stein EA, Yang Y. Group independent component analysis reveals consistent resting-state networks across multiple sessions. Brain Res. 2008 Nov 6; 1239:141-51.

Cheng, H., & Puce, A. (2014). Reducing respiratory effect in motion correction for EPI images with sequential slice acquisition order. *J Neurosci Methods, 227*, 83-89.

Chhatwal, J. P., & Sperling, R. A. (2012). Functional MRI of mnemonic networks across the spectrum of normal aging, mild cognitive impairment, and Alzheimer's disease. *J Alzheimers Dis, 31 Suppl 3*, S155-167.

Chia, J. M., Fischer, S. E., Wickline, S. A., & Lorenz, C. H. (2000). Performance of QRS detection for cardiac magnetic resonance imaging with a novel vectorcardiographic triggering method. *J Magn Reson Imaging, 12*(5), 678-688.

References

Chou, Y. H., Panych, L. P., Dickey, C. C., Petrella, J. R., & Chen, N. K. (2012). Investigation of long-term reproducibility of intrinsic connectivity network mapping: a resting-state fMRI study. *AJNR Am J Neuroradiol, 33*(5), 833-838.

Churchill, N.W., Oder, A., Abdi, H., Tam, F., Lee, W., Thomas, C., Ween, J.E., Graham, S.J., Strother, S.C. (2012). Optimizing preprocessing and analysis pipelines for single-subject fMRI. I. Standard temporal motion and physiological noise correction methods. *HumBrain Mapp*, 33(3), 609-27.

Churchill, N. W., & Strother, S. C. (2013). PHYCAA+: an optimized, adaptive procedure for measuring and controlling physiological noise in BOLD fMRI. *Neuroimage, 82*, 306-325.

Cordes, D., Haughton, V.M., Arfanakis, K., Carew, J.D., Turski, P.A., Moritz, C.H., Quigley, M.A., Meyerand, M.E. (2001). Frequencies contributing to functional connectivity in the cerebral cortex in "resting-state" data. *AJNR*, 22(7), 1326-33.

Cordes, D., Haughton, V.M., Arfanakis, K., Wendt, G.J., Turski, P.A., Moritz, C.H., Quigley, M.A., Meyerand, M.E. (2000). Mapping functionally related regions of brain with functional connectivity MR imaging. *AJNR*, 21(9), 1636-44

Cox, R. W. (1996). AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical Research, an International Journal, 29*, 162–173.

Craddock, R. C., Holtzheimer, P. E., 3rd, Hu, X. P., & Mayberg, H. S. (2009). Disease state prediction from resting state functional connectivity. *Magn Reson Med, 62*(6), 1619-1628.

Damoiseaux, J. S. (2012). Resting-state fMRI as a biomarker for Alzheimer's disease? *Alzheimers Res Ther, 4*(2), 8.

Damoiseaux, J.S., Beckmann, C.F., Arigita, E.J., Barkhof, F., Scheltens, P., Stam, C.J., Smith, S.M., Rombouts, S.A. (2008). Reduced resting-state brain activity in the "default network" in normal aging. Cereb Cortex, 18(8), 1856-64.

Damoiseaux, J. S., Rombouts, S. A., Barkhof, F., Scheltens, P., Stam, C. J., Smith, S. M., & Beckmann, C. F. (2006). Consistent resting-state networks across healthy subjects. *Proc Natl Acad Sci U S A, 103*(37), 13848-13853.

Davey, C. G., Harrison, B. J., Yucel, M., & Allen, N. B. (2012). Regionally specific alterations in functional connectivity of the anterior cingulate cortex in major depressive disorder. *Psychol Med, 42*(10), 2071-2081.

References

De Luca, M., Beckmann, C. F., De Stefano, N., Matthews, P. M., & Smith, S. M. (2006). fMRI resting state networks define distinct modes of long-distance interactions in the human brain. *Neuroimage, 29*(4), 1359-1367.

Desikan, R.S., Ségonne, F., Fischl, B., Quinn, B.T., Dickerson, B.C., Blacker, D., Buckner, R.L., Dale, A.M., Maguire, R.P., Hyman, B.T., Albert, M.S., Killiany, R.J. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage*, 31(3), 968-80.

Dosenbach, N.U., Nardos, B., Cohen, A.L., Fair, D.A., Power, J.D., Church, J.A., Nelson, S.M., Wig, G.S., Vogel, A.C., Lessov-Schlaggar, C.N., Barnes, K.A., Dubis, J.W., Feczko, E., Coalson, R.S., Pruett, J.R., Barch, D.M., Petersen, S.E., Schlaggar, B.L. (2010). Prediction of individual brain maturity using fMRI. *Science*, 329(5997), 1358-61

Drago, V., Babiloni, C., Bartrés-Faz, D., Caroli, A., Bosch, B., Hensch, T., Didic, M., Klafki, H.W., Pievani, M., Jovicich, J., Venturi, L., Spitzer, P., Vecchio, F., Schoenknecht, P., Wiltfang, J., Redolfi, A., Forloni, G., Blin, O., Irving, E., Davis, C., Hårdemark, H.G., Frisoni, G.B. (2011). Disease tracking markers for Alzheimer's disease at the prodromal (MCI) stage. *J Alzheimers Dis*, 26, Suppl 3, 159-99.

Du, Y., Allen, E. A., He, H., Sui, J., Wu, L., & Calhoun, V. D. (2016). Artifact removal in the context of group ICA: A comparison of single-subject and group approaches. *Hum Brain Mapp, 37*(3), 1005-1025.

Dunn, O. J. (1964). Multiple Comparisons Using Rank Sums. *Technometrics, 6*, 241-252.

Etkin, A., Prater, K. E., Schatzberg, A. F., Menon, V., and Greicius, M. D. (2009). Disrupted amygdalar subregion functional connectivity and evidence of a compensatory network in generalized anxiety disorder. *Arch. Gen. Psychiatry* 66, 1361–1372.

Fair, D.A., Cohen, A.L., Power, J.D., Dosenbach, N.U., Church, J.A., Miezin, F.M., Schlaggar, B.L., Petersen, S.E. (2009). Functional brain networks develop from a "local to distributed" organization. *PLoS Comput Biol*, 5(5):e1000381

Fair, D.A., Dosenbach, N.U., Church, J.A., Cohen, A.L., Brahmbhatt, S., Miezin, F.M., Barch, D.M., Raichle, M.E., Petersen, S.E., Schlaggar, B.L. (2007). Development of distinct control networks through segregation and integration. *Proc Natl Acad Sci USA*, 104(33), 13507-12.

Feinberg, D.A., Moeller, S., Smith, S.M., Auerbach, E., Ramanna, S., Gunther, M., Glasser, M.F., Miller, K.L., Ugurbil, K., Yacoub, E. (2010). Multiplexed echo planar imaging for sub-second whole brain FMRI and fast diffusion imaging. *PLoS One*, 5(12):e15710.

Feis, R.A., Smith, S.M., Filippini, N., Douaud, G., Dopper, E.G., Heise, V., Trachtenberg, A.J., van Swieten, J.C., van Buchem, M.A., Rombouts, S.A., Mackay, C.E. (2015). ICA-based artifact removal diminishes scan site differences in multi-center resting-state fMRI. *Front Neurosci*, 9:395.

Ferreira, L.K., Regina, A.C., Kovacevic, N., Martin, M.D., Santos, P.P., Carneiro, C.G., Kerr, D.S., Amaro, E., McIntosh, A.R., Busatto, G.F. (2015). Aging Effects on Whole-Brain Functional Connectivity in Adults Free of Cognitive and Psychiatric Disorders. *Cerebral cortex (New York, NY: 1991)*.

Fox, M. D., & Greicius, M. (2010). Clinical applications of resting state functional connectivity. *Frontiers in systems neuroscience*, *4*, 19.

Fox, M. D., & Raichle, M. E. (2007). Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Nat Rev Neurosci, 8*(9), 700-711.

Fox, M. D., Snyder, A. Z., Vincent, J. L., Corbetta, M., Van Essen, D. C., & Raichle, M. E. (2005). The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proceedings of the National Academy of Sciences of the United States of America, 102*(27), 9673–9678.

Fox, M. D., Zhang, D., Snyder, A. Z., & Raichle, M. E. (2009). The global signal and observed anticorrelated resting state brain networks. *J Neurophysiol, 101*(6), 3270-3283.

Franco, A. R., Mannell, M. V., Calhoun, V. D., & Mayer, A. R. (2013). Impact of analysis methods on the reproducibility and reliability of resting-state networks. *Brain Connect, 3*(4), 363-374.

Franco, A. R., Pritchard, A., Calhoun, V. D., & Mayer, A. R. (2009). Interrater and intermethod reliability of default mode network selection. *Human brain mapping*, *30*(7), 2293-2303.

Frank, R., & Hargreaves, R. (2003). Clinical biomarkers in drug discovery and development. *Nat Rev Drug Discov, 2*(7), 566-580.

Fransson, P. (2005). Spontaneous low-frequency BOLD signal fluctuations: an fMRI investigation of the resting-state default mode of brain function hypothesis. *Hum Brain Mapp, 26*(1), 15-29.

Fransson, P., & Marrelec, G. (2008). The precuneus/posterior cingulate cortex plays a pivotal role in the default mode network: Evidence from a partial correlation network analysis. *Neuroimage, 42*(3), 1178-1184.

Fraser, M. A., Shaw, M. E., & Cherbuin, N. (2015). A systematic review and meta-analysis of longitudinal hippocampal atrophy in healthy human ageing. *Neuroimage, 112*, 364-374.

Frisoni, G.B., Bocchetta, M., Chételat, G., Rabinovici, G.D., de Leon, M.J., Kaye, J., Reiman, E.M., Scheltens, P., Barkhof, F., Black, S.E., Brooks, D.J., Carrillo, M.C., Fox, N.C., Herholz, K., Nordberg, A., Jack, C.R., Jagust, W.J., Johnson, K.A., Rowe, C.C., Sperling, R.A., Thies, W., Wahlund, L.O., Weiner, M.W., Pasqualetti, P., Decarli, C.; ISTAART's NeuroImaging Professional Interest Area. (2013). Imaging markers for Alzheimer disease Which vs how. *Neurology*, *81*(5), 487-500.

Frisoni, G. B., & Visser, P. J. (2015). Biomarkers for Alzheimer's disease: a controversial topic. *Lancet Neurol, 14*(8), 781-783.

Friston, K. J. (2011). Functional and effective connectivity: a review. *Brain Connect, 1*(1), 13-36.

Galvin, J. E., Price, J. L., Yan, Z., Morris, J. C., & Sheline, Y. I. (2011). Resting bold fMRI differentiates dementia with Lewy bodies vs Alzheimer disease. *Neurology, 76*(21), 1797-1803.

Glover, G. H., Li, T. Q., & Ress, D. (2000). Image-based method for retrospective correction of physiological motion effects in fMRI: RETROICOR. *Magn Reson Med, 44*(1), 162-167.

Godenschweger, F., Kägebein, U., Stucht, D., Yarach, U., Sciarra, A., Yakupov, R., Lüsebrink, F., Schulze, P., Speck, O. (2016). Motion correction in MRI of the brain. *Physics in medicine and biology*, *61*(5), R32-56.

Greicius, M.D., Flores, B.H., Menon, V., Glover, G.H., Solvason, H.B., Kenna, H., Reiss, A.L., Schatzberg, A.F. (2007). Resting-state functional connectivity in major depression: abnormally increased contributions from subgenual cingulate cortex and thalamus. *Biological psychiatry*, *62*(5), 429-437.

Greicius, M. D., Krasnow, B., Reiss, A. L., & Menon, V. (2003). Functional connectivity in the resting brain: a network analysis of the default mode hypothesis. *Proc Natl Acad Sci U S A, 100*(1), 253-258.

Greicius, M. D., & Menon, V. (2004). Default-mode activity during a passive sensory task: uncoupled from deactivation but impacting activation. *J Cogn Neurosci, 16*(9), 1484-1492.

Greicius, M. D., Srivastava, G., Reiss, A. L., & Menon, V. (2004). Default-mode network activity distinguishes Alzheimer's disease from healthy aging: evidence from functional MRI. *Proc Natl Acad Sci U S A, 101*(13), 4637-4642.

Greicius, M. D., Barad, M., Ueno, T., and Mackey, S. C. (2008a). “Chronic pain remodels the brain’s salience network: a resting-state fMRI study,” in *14th International Meeting of the Organization for Human Brain* Mapping Melbourne, Australia.

Greve, D. N., & Fischl, B. (2009). Accurate and robust brain image alignment using boundary-based registration. *Neuroimage, 48*(1), 63-72.

Griffanti, L., Rolinski, M., Szewczyk-Krolikowski, K., Menke, R.A., Filippini, N., Zamboni, G., Jenkinson, M., Hu, M.T., Mackay, C.E. (2016). Challenges in the reproducibility of clinical studies with resting state fMRI: An example in early Parkinson's disease. *NeuroImage*, *124*, 704-713.

Griffanti, L., Salimi-Khorshidi, G., Beckmann, C.F., Auerbach, E.J., Douaud, G., Sexton, C.E., Zsoldos, E., Ebmeier, K.P., Filippini, N., Mackay, C.E., Moeller, S., Xu, J., Yacoub, E., Baselli, G., Ugurbil, K., Miller, K.L., Smith, SM. (2014). ICA-based artefact removal and accelerated fMRI acquisition for improved resting state network imaging. *Neuroimage*, 95, 232-47.

Guo, C. C., Kurth, F., Zhou, J., Mayer, E. A., Eickhoff, S. B., Kramer, J. H., & Seeley, W. W. (2012). One-year test-retest reliability of intrinsic connectivity network fMRI in older adults. *Neuroimage, 61*(4), 1471-1483.

Gusnard, D. A., Raichle, M. E., & Raichle, M. E. (2001). Searching for a baseline: functional imaging and the resting human brain. *Nat Rev Neurosci, 2*(10), 685-694.

Habeck, C., Risacher, S., Lee, G.J., Glymour, M.M., Mormino, E., Mukherjee, S., Kim, S., Nho, K., DeCarli, C., Saykin, A.J., Crane, P.K. Alzheimer’s Disease Neuroimaging Initiative. (2012). Relationship between baseline brain metabolism measured using [^{18}F]FDG PET and memory and executive function in prodromal and early Alzheimer's disease. *Brain Imaging Behav*, 6(4), 568-83.

Hafkemeijer, A., van der Grond, J., & Rombouts, S. A. (2012). Imaging the default mode network in aging and dementia. *Biochim Biophys Acta, 1822*(3), 431-441.

Hale, J. R., Brookes, M. J., Hall, E. L., Zumer, J. M., Stevenson, C. M., Francis, S. T., & Morris, P. G. (2010). Comparison of functional connectivity in default mode and sensorimotor networks at 3 and 7T. *MAGMA, 23*(5-6), 339-349.

Hallquist, M. N., Hwang, K., & Luna, B. (2013). The nuisance of nuisance regression: spectral misspecification in a common approach to resting-state fMRI preprocessing reintroduces noise and obscures functional connectivity. *Neuroimage, 82*, 208-225.

Hampel, H., Wilcock, G., Andrieu, S., Aisen, P., Blennow, K., Broich, K., Carrillo, M., Fox, N.C., Frisoni, G.B., Isaac, M., Lovestone, S., Nordberg, A., Prvulovic, D., Sampaio, C., Scheltens, P., Weiner, M., Winblad, B., Coley, N., Vellas, B. (2011). Oxford Task Force Group. Biomarkers for Alzheimer's disease therapeutic trials. Prog Neurobiol, 95(4), 579-93.

Harvey, A.K., Pattinson, K.T., Brooks, J.C., Mayhew, S.D., Jenkinson, M., Wise, R.G. (2008). Brainstem functional magnetic resonance imaging: disentangling signal from physiological noise. J Magn Reson Imaging, 28(6), 1337-44.

Hawellek, D. J., Hipp, J. F., Lewis, C. M., Corbetta, M., & Engel, A. K. (2011). Increased functional connectivity indicates the severity of cognitive impairment in multiple sclerosis. *Proc Natl Acad Sci U S A, 108*(47), 19066-19071.

Hedden, T., Van Dijk, K. R., Becker, J. A., Mehta, A., Sperling, R. A., Johnson, K. A., & Buckner, R. L. (2009). Disruption of functional connectivity in clinically normal older adults harboring amyloid burden. *J Neurosci, 29*(40), 12686-12694.

Hodkinson, D.J., O'Daly, O., Zunszain, P.A., Pariante, C.M., Lazurenko, V., Zelaya, F.O., Howard, M.A., Williams, S.C. (2014). Circadian and homeostatic modulation of functional connectivity and regional cerebral blood flow in humans under normal entrained conditions. *J Cereb Blood Flow Metab*, 34(9), 1493-9.

Horn, A., Ostwald, D., Reisert, M., & Blankenburg, F. (2014). The structural-functional connectome and the default mode network of the human brain. *Neuroimage, 102 Pt 1*, 142-151.

Horovitz, S. G., Braun, A. R., Carr, W. S., Picchioni, D., Balkin, T. J., Fukunaga, M., & Duyn, J. H. (2009). Decoupling of the brain's default mode network during deep sleep. *Proc Natl Acad Sci U S A, 106*(27), 11376-11381.

Horwitz, B., Warner, B., Fitzer, J., Tagamets, M. A., Husain, F. T., & Long, T. W. (2005). Investigating the neural basis for functional and effective connectivity. Application to fMRI. *Philos Trans R Soc Lond B Biol Sci, 360*(1457), 1093-1108.

Hosseini, S. M., & Kesler, S. R. (2013). Comparing connectivity pattern and small-world organization between structural correlation and resting-state networks in healthy adults. *Neuroimage, 78*, 402-414.

Hu, X., Le, T. H., Parrish, T., & Erhard, P. (1995). Retrospective estimation and correction of physiological fluctuation in functional MRI. *Magnetic resonance in medicine*, *34*(2), 201-212.

References

Huang, C.C., Hsieh, W.J., Lee, P.L., Peng, L.N., Liu, L.K., Lee, W.J., Huang, J.K., Chen, L.K., Lin, C.P. (2015). Age-related changes in resting-state networks of a large sample size of healthy elderly. *CNS Neurosci Ther*, 21(10), 817-25.

Huettel, S. A., Song, A.W., & McCarthy, G. (2008). Functional Magnetic Resonance Imaging (2nd edition). *Sinauer Associates, Inc: Sunderland, Massachusetts, USA*.

Hutton, C., Josephs, O., Stadler, J., Featherstone, E., Reid, A., Speck, O., Bernarding, J., Weiskopf, N. (2011). The impact of physiological noise correction on fMRI at 7 T. *Neuroimage*, 57(1), 101-12.

Jack, C.R., Knopman, D.S., Weigand, S.D., Wiste, H.J., Vemuri, P., Lowe, V., Kantarci, K., Gunter, J.L., Senjem, M.L., Ivnik, R.J., Roberts, R.O., Rocca, W.A., Boeve, B.F., Petersen, R.C. (2012). An operational approach to National Institute on Aging-Alzheimer's Association criteria for preclinical Alzheimer disease. *Ann Neurol,* 71(6), 765-7.

Kong, J., Gollub, R. L., Webb, J. M., Kong, J. T., Vangel, M. G., & Kwong, K. (2007). Test–retest study of fMRI signal change evoked by electroacupuncture stimulation. *Neuroimage,* 34(3), 1171-1181.

Thada, V., & Jaglan, V. (2013). Comparison of Jaccard, Dice, Cosine Similarity Coefficient To Find Best Fitness Value for Web Retrieved Documents Using Genetic Algorithm. *International Journal of Innovations in Engineering and Technology, 2*(4), 202-205.

Jao, T., Li, C.W., Vértes, P.E., Wu, C.W., Achard, S., Hsieh, C.H., Liou, C.H., Chen, J.H., Bullmore, E.T. (2016). Large-Scale Functional Brain Network Reorganization During Taoist Meditation. *Brain Connect*, 6(1), 9-24.

Jenkinson, M., Bannister, P., Brady, J. M. and Smith, S. M. (2002). Improved Optimisation for the Robust and Accurate Linear Registration and Motion Correction of Brain Images. *Neuroimage, 17*(2), 825-841.

Jenkinson, M., Beckmann, C. F., Behrens, T. E., Woolrich, M. W., & Smith, S. M. (2012). Fsl. *Neuroimage, 62*(2), 782-790.

Jenkinson, M., & Smith, S. (2001). A global optimisation method for robust affine registration of brain images. *Med Image Anal, 5*(2), 143-156.

Jeong, B., Choi, J., & Kim, J. W. (2012). MRI study on the functional and spatial consistency of resting state-related independent components of the brain network. *Korean J Radiol, 13*(3), 265-274.

Jo, H. J., Saad, Z. S., Simmons, W. K., Milbury, L. A., & Cox, R. W. (2010). Mapping sources of correlation in resting state FMRI, with artifact detection and removal. *Neuroimage, 52*(2), 571-582.

Jones, T. B., Bandettini, P. A., & Birn, R. M. (2008). Integration of motion correction and physiological noise regression in fMRI. *Neuroimage, 42*(2), 582-590.

Jovicich, J., Marizzoni, M., Bosch, B., Bartrés-Faz, D., Arnold, J., Benninghoff, J., Wiltfang, J., Roccatagliata, L., Picco, A., Nobili, F., Blin, O., Bombois, S., Lopes, R., Bordet, R., Chanoine, V., Ranjeva, J.P., Didic, M., Gros-Dagnac, H., Payoux, P., Zoccatelli, G., Alessandrini, F., Beltramello, A., Bargalló, N., Ferretti, A., Caulo, M., Aiello, M., Ragucci, M., Soricelli, A., Salvadori, N., Tarducci, R., Floridi, P., Tsolaki, M., Constantinidis, M., Drevelegas, A., Rossini, P.M., Marra, C., Otto, J., Reiss-Zimmermann, M., Hoffmann, K.T., Galluzzi, S., Frisoni, G.B. (2014). PharmaCog Consortium. Multisite longitudinal reliability of tract-based spatial statistics in diffusion tensor imaging of healthy elderly subjects. *Neuroimage*, 101, 390-403.

Jovicich, J., Marizzoni, M., Sala-Llonch, R., Bosch, B., Bartrés-Faz, D., Arnold, J., Benninghoff, J., Wiltfang, J., Roccatagliata, L., Nobili, F., Hensch, T., Tränkner, A., Schönknecht, P., Leroy, M., Lopes, R., Bordet, R., Chanoine, V., Ranjeva, J.P., Didic, M., Gros-Dagnac, H., Payoux, P., Zoccatelli, G., Alessandrini, F., Beltramello, A., Bargalló, N., Blin, O., Frisoni, G.B. (2013). PharmaCog Consortium. Brain morphometry reproducibility in multi-center 3T MRI studies: a comparison of cross-sectional and longitudinal segmentations. *Neuroimage*, 83, 472-84.

Jovicich, J., Minati, L., Marizzoni, M., Marchitelli, R., Sala-Llonch, R., Bartrés-Faz, D., Arnold, J., Benninghoff, J., Fiedler, U., Roccatagliata, L., Picco, A., Nobili, F., Blin, O., Bombois, S., Lopes, R,, Bordet, R., Sein, J., Ranjeva, J.P., Didic, M., Gros-Dagnac, H., Payoux, P., Zoccatelli, G., Alessandrini, F., Beltramello, A., Bargalló, N., Ferretti, A., Caulo, M., Aiello, M., Cavaliere, C., Soricelli, A., Parnetti, L., Tarducci, R., Floridi, P., Tsolaki, M., Constantinidis, M., Drevelegas, A., Rossini, P.M., Marra, C., Schönknecht, P., Hensch, T., Hoffmann, K.T., Kuijer, J.P., Visser, P.J., Barkhof, F., Frisoni, G.B. (2016). The PharmaCog Consortium. Longitudinal reproducibility of default-mode network connectivity in healthy elderly participants: A multicentric resting-state fMRI study. *Neuroimage*, 124 (Pt A), 442-54.

Jovicich, J., Marizzoni, M., Sala-Llonch, R., Bosch, B., Bartrés-Faz, D., Arnold, J., Benninghoff, J., Wiltfang, J., Roccatagliata, L., Nobili, F., Hensch, T., Tränkner, A., Schönknecht, P., Leroy, M., Lopes, R., Bordet, R., Chanoine, V., Ranjeva, J.P., Didic, M., Gros-Dagnac, H., Payoux, P., Zoccatelli, G., Alessandrini, F., Beltramello, A., Bargalló, N., Blin, O., Frisoni, G.B, The PharmaCog Consortium. (2013). Brain morphometry reproducibility in multi-center 3T MRI studies: a comparison of cross-sectional and longitudinal segmentations. *Neuroimage, 83*, 472-484.

Katz, R. (2004a). Biomarkers and surrogate markers: an FDA perspective. *NeuroRx, 1*(2), 189-195.

Katz, R. (2004b). FDA: evidentiary standards for drug development and approval. *NeuroRx, 1*(3), 307-316.

Kelly, A.M., Uddin L.Q., Biswal B.B., Castellanos F.X., Milham M.P. (2008). Competition between functional brain networks mediates behavioral variability. *Neuroimage*, *39*(1), 527-537.

Kenet, T., Bibitchkov, D., Tsodyks, M., Grinvald, A., & Arieli, A. (2003). Spontaneously emerging cortical representations of visual attributes. *Nature, 425*(6961), 954-956.

Kim, B., Yeo, D. T., & Bhagalia, R. (2008). Comprehensive mathematical simulation of functional magnetic resonance imaging time series including motion-related image distortion and spin saturation effect. *Magn Reson Imaging, 26*(2), 147-159.

Kiviniemi, V., Kantola, J. H., Jauhiainen, J., Hyvarinen, A., & Tervonen, O. (2003). Independent component analysis of nondeterministic fMRI signal sources. *Neuroimage, 19*(2 Pt 1), 253-260.

Koch, W., Teipel, S., Mueller, S., Benninghoff, J., Wagner, M., Bokde, A.L., Hampel, H., Coates, U., Reiser, M., Meindl, T. (2012). Diagnostic power of default mode network resting state fMRI in the detection of Alzheimer's disease. *Neurobiol Aging*, 33(3), 466-78.

Laufs, H., Krakow, K., Sterzer, P., Eger, E., Beyerle, A., Salek-Haddadi, A., & Kleinschmidt, A. (2003). Electroencephalographic signatures of attentional and cognitive default modes in spontaneous brain activity fluctuations at rest. *Proc Natl Acad Sci USA, 100*(19), 11053-11058.

Lee, M. H., Smyser, C. D., & Shimony, J. S. (2013). Resting-state fMRI: a review of methods and clinical applications. *AJNR Am J Neuroradiol, 34*(10), 1866-1872.

Leech, R., Braga, R., & Sharp, D. J. (2012). Echoes of the brain within the posterior cingulate cortex. *JNeurosci, 32*(1), 215-222.

Leech, R., Kamourieh, S., Beckmann, C. F., & Sharp, D. J. (2011). Fractionating the default mode network: distinct contributions of the ventral and dorsal posterior cingulate cortex to cognitive control. *J Neurosci, 31*(9), 3217-3224.

Leopold, D. A., & Logothetis, N. K. (2003). Spatial patterns of spontaneous local field activity in the monkey visual cortex. *Rev Neurosci, 14*(1-2), 195-205.

Liao, W. M., D; Zhang, Z; Pan, Z; Ding, J; Gong, Q; Yang, Y; Chen, H;. (2009). Evaluating the effective connectivity of resting state networks using conditional Granger causality. *BIOLOGICAL CYBERNETICS, 102*(1), 57-69.

Li, K., Guo, L., Nie, J., Li, G., & Liu, T. (2009). Review of methods for functional brain connectivity detection using fMRI. *Computerized Medical Imaging and Graphics*, *33*(2), 131-139.

Liao, X.H., Xia, M. R., Xu, T., Dai, Z.J., Cao, X.Y., Niu, H.J., Zuo, X.N., Zang, Y.F., He, Y. (2013). Functional brain hubs and their test–retest reliability: a multiband resting-state functional MRI study. *Neuroimage*, *83*, 969-982.

Liu, C.S., Miki, A., Hulvershorn, J., Bloy, L., Gualtieri, E.E., Liu, G.T., Leigh, J.S., Haselgrove, J.C., Elliott, M.A. (2006). Spatial and temporal characteristics of physiological noise in fMRI at 3T. *Acad Radiol*, 13(3), 313-23.

Lowe, M. J. (2012). The emergence of doing "nothing" as a viable paradigm design. *Neuroimage, 62*(2), 1146-1151.

Lowe, M. J., Dzemidzic, M., Lurito, J. T., Mathews, V. P., & Phillips, M. D. (2000). Correlations in low-frequency BOLD fluctuations reflect cortico-cortical connections. *Neuroimage, 12*(5), 582-587.

Lowe, M. J., Mock, B. J., & Sorenson, J. A. (1998). Functional connectivity in single and multislice echoplanar imaging using resting-state fluctuations. *Neuroimage, 7*(2), 119-132.

Lund, T. E. (2001). fcMRI--mapping functional connectivity or correlating cardiac-induced noise? *Magnetic Resonance in Medicine: Official Journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine, 46*(3), 628–629.

Lutkenhoff, E. S., Rosenberg, M., Chiang, J., Zhang, K., Pickard, J. D., Owen, A. M., & Monti, M. M. (2014). Optimized brain extraction for pathological brains (optiBET). *PLoS One, 9*(12), e115551.

Maitra, R. (2010). A re-defined and generalized percent-overlap-of-activation measure for studies of fMRI reproducibility and its use in identifying outlier activation maps. *Neuroimage, 50*(1), 124-135.

Marchitelli, R., Minati, L., Marizzoni, M., Bosch, B., Bartrés−Faz, D., Müller, B. W., Wiltfang, J., Fiedler, U., Roccatagliata, L., Picco, A., Nobili, F., Blin, O., Bombois, S., Lopes, R., Bordet, R., Sein, J., Ranjeva, J. P., Didic, M., Gros-Dagnac, H., Payoux, P., Zoccatelli, G., Alessandrini, F., Beltramello, A., Bargalló, N., Ferretti, A., Caulo, M., Aiello, M., Cavaliere, C., Soricelli, A., Parnetti, L., Tarducci, R., Floridi, P., Tsolaki, M., Constantinidis, M., Drevelegas, A., Rossini, P. M., Marra, C., Schönknecht, P., Hensch, T., Hoffmann, K. T., Kuijer, J. P., Visser, P. J., Barkhof, F., Frisoni, G. B., Jovicich, J. (2016). Test−retest reliability of the default mode network in a multi−centric fMRI study of healthy elderly: Effects of data−driven physiological noise correction techniques. *Human Brain Mapping*.

Marstaller, L., Williams, M., Rich, A., Savage, G., & Burianova, H. (2015). Aging and large-scale functional networks: white matter integrity, gray matter volume, and functional connectivity in the resting state. *Neuroscience, 290*, 369-378.

Mat Roni, S. (2014). Introduction to SPSS, School of Business. *Edith Cowan University, Australia.*

Mazoyer, B., Zago, L., Mellet, E., Bricogne, S., Etard, O., Houdé, O., Crivello, F., Joliot, M., Petit, L., Tzourio-Mazoyer, N. (2001). Cortical networks for working memory and executive functions sustain the conscious resting state in man. *Brain Res Bull*, 54(3), 287-98.

McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological methods*, *1*(1), 30.

McGuire, P. K., Paulesu, E., Frackowiak, R. S., & Frith, C. D. (1996). Brain activity during stimulus independent thought. *Neuroreport, 7*(13), 2095-2099.

McKeown, M. J., Makeig, S., Brown, G. G., Jung, T. P., Kindermann, S. S., Bell, A. J., & Sejnowski, T. J. (1997). *Analysis of fMRI data by blind separation into independent spatial components* (No. NHRC-REPT-97-42). NAVAL HEALTH RESEARCH CENTER SAN DIEGO CA.

Meier, T.B., Desphande, A.S., Vergun, S., Nair, V.A., Song, J., Biswal, B.B., Meyerand, M.E., Birn, R.M., Prabhakaran, V. (2012). Support vector machine classification and characterization of age-related reorganization of functional brain networks. *Neuroimage*, 60(1), 601-13.

Meindl, T., Teipel, S., Elmouden, R., Mueller, S., Koch, W., Dietrich, O., Coates, U., Reiser, M., Glaser, C. (2010). Test-retest reproducibility of the default-mode network in healthy individuals. *Hum Brain Mapp*, 31(2), 237-46.

Molloy, E. K., Meyerand, M. E., & Birn, R. M. (2014). The influence of spatial resolution and smoothing on the detectability of resting-state and task fMRI. *Neuroimage, 86*, 221-230.

Müller, R., & Büttner, P. (1994). A critical discussion of intraclass correlation coefficients. *Statistics in medicine*, *13*(23−24), 2465-2476.

Mueller, S., Wang, D., Fox, M.D., Pan, R., Lu, J., Li, K., Sun, W., Buckner, R.L., Liu, H. (2015). Reliability correction for functional connectivity: Theory and implementation. *Hum Brain Mapp*, 36(11), 4664-80.

Mulders, P. C., van Eijndhoven, P. F., Schene, A. H., Beckmann, C. F., & Tendolkar, I. (2015). Resting-state functional connectivity in major depressive disorder: A review. *Neurosci Biobehav Rev, 56*, 330-344.

Murphy, K., Birn, R. M., & Bandettini, P. A. (2013). Resting-state fMRI confounds and cleanup. *Neuroimage, 80*, 349-359.

Murphy, K., Birn, R. M., Handwerker, D. A., Jones, T. B., & Bandettini, P. A. (2009). The impact of global signal regression on resting state correlations: are anti-correlated networks introduced? *Neuroimage, 44*(3), 893-905.

Nicolini, P., Ciulla, M. M., De Asmundis, C., Magrini, F., & Brugada, P. (2012). The prognostic value of heart rate variability in the elderly, changing the perspective: from sympathovagal balance to chaos theory. *Pacing Clin Electrophysiol, 35*(5), 622-638.

Nir, Y., Mukamel, R., Dinstein, I., Privman, E., Harel, M., Fisch, L., Gelbard-Sagiv, H., Kipervasser, S., Andelman, F., Neufeld, M.Y., Kramer, U., Arieli, A., Fried, I., Malach, R. (2008). Interhemispheric correlations of slow spontaneous neuronal fluctuations revealed in human sensory cortex. *Nat Neurosci*, 11(9), 1100-8.

Noseworthy, M. D., Alfonsi, J., & Bells, S. (2003). Attenuation of brain BOLD response following lipid ingestion. *Hum Brain Mapp, 20*(2), 116-121.

O'Connor, A. (2010). The Physical Reality of fMRI Slice Acquisition Order. Retrieved from http://akiraoconnor.org/2010/09/01/the-physical-reality-of-fmri-slice-acquisition-order

Ogawa, S., Lee, T. M., Kay, A. R., & Tank, D. W. (1990). Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proc Natl Acad Sci U S A, 87*(24), 9868-9872.

Ogawa, S., Tank, D. W., Menon, R., Ellermann, J. M., Kim, S. G., Merkle, H., & Ugurbil, K. (1992). Intrinsic signal changes accompanying sensory stimulation: functional brain mapping with magnetic resonance imaging. *Proc Natl Acad Sci U S A, 89*(13), 5951-5955.

Palop, J. J., Chin, J., & Mucke, L. (2006). A network dysfunction perspective on neurodegenerative diseases. *Nature, 443*(7113), 768-773.

Pan, W. J., Thompson, G. J., Magnuson, M. E., Jaeger, D., & Keilholz, S. (2013). Infraslow LFP correlates to resting-state fMRI BOLD signals. *Neuroimage, 74*, 288-297.

Parker, D., Rotival, G., Laine, A., & Razlighi, Q. R. (2014). Retrospective Detection of Interleaved Slice Acquisition Parameters from Fmri Data. *Proc IEEE Int Symp Biomed Imaging, 2014*, 37-40.

Patriat, R., Molloy, E.K., Meier, T.B., Kirk, G.R., Nair, V.A., Meyerand, M.E., Prabhakaran, V., Birn, R.M. (2013). The effect of resting condition on resting-state fMRI reliability and consistency: a comparison between resting with eyes open, closed, and fixated. *Neuroimage*, 78, 463-73.

Petrella, J. R., Sheldon, F. C., Prince, S. E., Calhoun, V. D., & Doraiswamy, P. M. (2011). Default mode network connectivity in stable vs progressive mild cognitive impairment. *Neurology, 76*(6), 511-517.

Poldrack, R. A., Mumford, J. A., & Nichols, T. E. (2011). *Handbook of functional MRI data analysis*. Cambridge University Press.

Poncelet, B. P., Wedeen, V. J., Weisskoff, R. M., & Cohen, M. S. (1992). Brain parenchyma motion: measurement with cine echo-planar MR imaging. *Radiology, 185*(3), 645-651.

Power, J. D., Barnes, K. A., Snyder, A. Z., Schlaggar, B. L., & Petersen, S. E. (2012). Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *Neuroimage, 59*(3), 2142-2154.

Power, J. D., Mitra, A., Laumann, T. O., Snyder, A. Z., Schlaggar, B. L., & Petersen, S. E. (2014). Methods to detect, characterize, and remove motion artifact in resting state fMRI. *Neuroimage, 84*, 320-341.

Power, J. D., Schlaggar, B. L., & Petersen, S. E. (2015). Recent progress and outstanding issues in motion correction in resting state fMRI. *Neuroimage, 105*, 536-551.

Prestia, A., Caroli, A., van der Flier, W.M., Ossenkoppele, R., Van Berckel, B., Barkhof, F., Teunissen, C.E., Wall, A.E., Carter, S.F., Schöll, M., Choo, I.H., Nordberg, A., Scheltens, P.,

Frisoni, G.B. (2013). Prediction of dementia in MCI patients based on core diagnostic markers for Alzheimer disease. *Neurology,* 80(11), 1048-56.

Pruim, R. H., Mennes, M., Buitelaar, J. K., & Beckmann, C. F. (2015a). Evaluation of ICA-AROMA and alternative strategies for motion artifact removal in resting state fMRI. *Neuroimage, 112*, 278-287.

Pruim, R. H., Mennes, M., van Rooij, D., Llera, A., Buitelaar, J. K., & Beckmann, C. F. (2015b). ICA-AROMA: A robust ICA-based strategy for removing motion artifacts from fMRI data. *Neuroimage, 112*, 267-277.

Rack-Gomer, A. L., Liau, J., & Liu, T. T. (2009). Caffeine reduces resting-state BOLD functional connectivity in the motor cortex. *Neuroimage, 46*(1), 56-63.

Richiardi, J., Altmann, A., Milazzo, A.C., Chang, C., Chakravarty, M.M., Banaschewski, T., Barker, G.J., Bokde, A.L., Bromberg, U., Büchel, C., Conrod, P., Fauth-Bühler, M., Flor, H., Frouin, V., Gallinat, J., Garavan, H., Gowland, P., Heinz, A., Lemaître, H., Mann, K.F., Martinot, J.L., Nees, F., Paus, T., Pausova, Z., Rietschel, M., Robbins, T.W., Smolka, M.N., Spanagel, R., Ströhle, A., Schumann, G., Hawrylycz, M., Poline, J.B., Greicius, M.D. (2015). IMAGEN consortium. BRAIN NETWORKS. Correlated gene expression supports synchronous activity in brain networks. *Science*, 348(6240), 1241-4.

Richter, W. S. (2006). Imaging biomarkers as surrogate endpoints for drug development. *Eur J Nucl Med Mol Imaging, 33 Suppl 1*, 6-10.

Roche, A. (2011). A four-dimensional registration algorithm with application to joint correction of motion and slice timing in fMRI. *IEEE Trans Med Imaging, 30*(8), 1546-1554.

Rombouts, S. A., Barkhof, F., Hoogenraad, F. G., Sprenger, M., Valk, J., & Scheltens, P. (1997). Test-retest analysis with functional MR of the activated area in the human visual cortex. *AJNR Am J Neuroradiol, 18*(7), 1317-1322.

Rombouts, S. A., Barkhof, F., Hoogenraad, F. G., Sprenger, M., & Scheltens, P. (1998). Within-subject reproducibility of visual activation patterns with functional magnetic resonance imaging using multislice echo planar imaging. *Magnetic Resonance Imaging, 16*, 105–113.

Rosazza, C., & Minati, L. (2011). Resting-state brain networks: literature review and clinical applications. *Neurol Sci, 32*(5), 773-785.

Rosazza, C., Minati, L., Ghielmetti, F., Mandelli, M. L., & Bruzzone, M. G. (2012). Functional connectivity during resting-state functional MR imaging: study of the correspondence

between independent component analysis and region-of-interest-based methods. *AJNR Am J Neuroradiol, 33*, 180–187.

Rousson, V., Gasser, T., & Seifert, B. (2002). Assessing intrarater, interrater and test–retest reliability of continuous measurements. *Statistics in medicine*, *21*(22), 3431-3446.

Sala-Llonch, R., Pena-Gomez, C., Arenaza-Urquijo, E. M., Vidal-Pineiro, D., Bargallo, N., Junque, C., & Bartres-Faz, D. (2012). Brain connectivity during resting state and subsequent working memory task predicts behavioural performance. *Cortex, 48*(9), 1187-1196.

Salimi-Khorshidi, G., Douaud, G., Beckmann, C. F., Glasser, M. F., Griffanti, L., & Smith, S. M. (2014). Automatic denoising of functional MRI data: Combining independent component analysis and hierarchical fusion of classifiers. *Neuroimage, 90*, 449–468.

Salvador, R., Suckling, J., Coleman, M. R., Pickard, J. D., Menon, D., & Bullmore, E. (2005). Neurophysiological architecture of functional magnetic resonance images of human brain. *Cereb Cortex, 15*(9), 1332-1342.

Sambataro, F., Murty, V. P., Callicott, J. H., Tan, H. Y., Das, S., Weinberger, D. R., & Mattay, V. S. (2010). Age-related alterations in default mode network: impact on working memory performance. *Neurobiol Aging, 31*(5), 839-852.

Santelli, C., Nezafat, R., Goddu, B., Manning, W. J., Smink, J., Kozerke, S., & Peters, D. C. (2011). Respiratory bellows revisited for motion compensation: preliminary experience for cardiovascular MR. *Magn Reson Med, 65*(4), 1097-1102.

Sapsford, K. E., Tezak, Z., Kondratovich, M., Pacanowski, M. A., Zineh, I., & Mansfield, E. (2010). Biomarkers to improve the benefit/risk balance for approved therapeutics: a US FDA perspective on personalized medicine. *Ther Deliv, 1*(5), 631-641.

Sarkka, S., Solin, A., Nummenmaa, A., Vehtari, A., Auranen, T., Vanni, S., & Lin, F. H. (2012). Dynamic retrospective filtering of physiological noise in BOLD fMRI: DRIFTER. *Neuroimage, 60*(2), 1517-1527.

Satterthwaite, T.D., Elliott, M.A., Gerraty, R.T., Ruparel, K., Loughead, J., Calkins, M.E., Eickhoff, S.B., Hakonarson, H., Gur, R.C., Gur, R.E., Wolf, D.H. (2013). An improved framework for confound regression and filtering for control of motion artifact in the preprocessing of resting-state functional connectivity data. *Neuroimage*, 64, 240-56.

Satterthwaite, T.D., Wolf, D.H., Loughead, J., Ruparel, K., Elliott, M.A., Hakonarson, H., Gur, R.C., Gur, R.E. (2012). Impact of in-scanner head motion on multiple measures offunctional

connectivity: relevance for studies of neurodevelopment in youth. *Neuroimage*, 60(1), 623-32.

Saumier, D., Duong, A., Haine, D., Garceau, D., & Sampalis, J. (2009). Domain-specific cognitive effects of tramiprosate in patients with mild to moderate Alzheimer's disease: ADAS-cog subscale results from the Alphase Study. *J Nutr Health Aging, 13*(9), 808-812.

Schopf, V., Kasess, C. H., Lanzenberger, R., Fischmeister, F., Windischberger, C., & Moser, E. (2010). Fully exploratory network ICA (FENICA) on resting-state fMRI data. *J Neurosci Methods, 192*(2), 207-213.

Schöpf, V., Windischberger, C., Robinson, S., Kasess, C.H., Fischmeister, F.P., Lanzenberger, R., Albrecht, J., Kleemann, A.M., Kopietz, R., Wiesmann, M., Moser, E. (2011). Model-free fMRI group analysis using FENICA. *Neuroimage*, 55(1), 185-93.

Schulz, S., Adochiei, F. C., Edu, I. R., Schroeder, R., Costin, H., Bar, K. J., & Voss, A. (2013). Cardiovascular and cardiorespiratory coupling analyses: a review. *Philos Trans A Math Phys Eng Sci, 371*(1997), 20120191.

Seeley, W. W., Allman, J. M., Carlin, D. A., Crawford, R. K., Macedo, M. N., Greicius, M. D., Dearmond, S. J., and Miller, B. L. (2007). Divergent social functioning in behavioral variant frontotemporal dementia and Alzheimer disease: reciprocal networks and neuronal evolution. *Alzheimer Dis. Assoc. Disord.,* 21, S50–57.

Seeley, W. W., Crawford, R. K., Miller, B. L., and Greicius, M. D. (2008). "Cortical neurodegeneration syndromes target human structural-functional covariance networks," in *14th International Meeting of the Organization for Human Brain Mapping Melbourne*, Australia.

Shannon, B. J., & Buckner, R. L. (2004). Functional-anatomic correlates of memory retrieval that suggest nontraditional processing roles for multiple distinct regions within posterior parietal cortex. *J Neurosci, 24*(45), 10084-10092.

Shehzad, Z., Kelly, A.M., Reiss, P.T., Gee, D.G., Gotimer, K., Uddin, L.Q., Lee, S.H., Margulies, D.S., Roy, A.K., Biswal, B.B., Petkova, E., Castellanos, F.X., Milham, M.P. (2009). The resting brain: unconstrained yet reliable. Cereb Cortex, 19(10), 2209-29.

Sheline, Y. I., Barch, D. M., Price, J. L., Rundle, M. M., Vaishnavi, S. N., Snyder, A. Z., Mintun, M. A., Wang, S., Coalson, R. S., Raichle, M. E. (2009). The default mode network and self-referential processes in depression. *Proc Natl Acad Sci USA*, 106(6), 1942-7.

Sheline, Y. I., Raichle, M. E., Snyder, A. Z., Morris, J. C., Head, D., Wang, S., & Mintun, M. A. (2010). Amyloid plaques disrupt resting state default mode network connectivity in cognitively normal elderly. *Biol Psychiatry, 67*(6), 584-587.

Shen, X., Tokoglu, F., Papademetris, X., & Constable, R. T. (2013). Groupwise whole-brain parcellation from resting-state fMRI data for network node identification. *Neuroimage, 82*, 403-415.

Shirer, W. R., Ryali, S., Rykhlevskaia, E., Menon, V., & Greicius, M. D. (2012). Decoding subject-driven cognitive states with whole-brain connectivity patterns. *Cereb Cortex, 22*(1), 158-165.

Shmuel, A., & Leopold, D. A. (2008). Neuronal correlates of spontaneous fluctuations in fMRI signals in monkey visual cortex: Implications for functional connectivity at rest. *Hum Brain Mapp, 29*(7), 751-761.

Shmuel, A., Yacoub, E., Pfeuffer, J., Van de Moortele, P. F., Adriany, G., Hu, X., & Ugurbil, K. (2002). Sustained negative BOLD, blood flow and oxygen consumption response and its coupling to the positive response in the human brain. *Neuron, 36*(6), 1195-1210.

Shmueli, K., van Gelderen, P., de Zwart, J. A., Horovitz, S. G., Fukunaga, M., Jansma, J. M., & Duyn, J. H. (2007). Low-frequency fluctuations in the cardiac rate as a source of variance in the resting-state fMRI BOLD signal. *Neuroimage, 38*(2), 306-320.

Shoji, M. (2012). Biomarker of dementia: clinical and preclinical diagnosis. *Nihon Rinsho, 70*(5), 857-863.

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: uses in assessing rater reliability. *Psychol Bull, 86*(2), 420-428.

Shulman, G. L., Corbetta, M., Buckner, R. L., Fiez, J. A., Miezin, F. M., Raichle, M. E., & Petersen, S. E. (1997). Common Blood Flow Changes across Visual Tasks: I. Increases in Subcortical Structures and Cerebellum but Not in Nonvisual Cortex. *J Cogn Neurosci, 9*(5), 624-647.

Singh, K. D., & Fawcett, I. P. (2008). Transient and linearly graded deactivation of the human default-mode network by a visual detection task. *Neuroimage, 41*, 100–112.

Sladky, R., Friston, K. J., Trostl, J., Cunnington, R., Moser, E., & Windischberger, C. (2011). Slice-timing effects and their correction in functional MRI. *Neuroimage, 58*(2), 588-594.

Smith, S.M., Miller, K.L., Moeller, S., Xu, J., Auerbach, E.J., Woolrich, M.W., Beckmann, C.F., Jenkinson, M., Andersson, J., Glasser, M.F., Van Essen, D.C., Feinberg, D.A., Yacoub, E.S.,

Ugurbil, K. (2012). Temporally-independent functional modes of spontaneous brain activity. Proc Natl Acad Sci USA, 109(8), 3131-6.

Somandepalli, K., Kelly, C., Reiss, P. T., Zuo, X. N., Craddock, R. C., Yan, C. G., Petkova, E., Castellanos, F. X., Milham, M. P., Di Martino, A. (2015). Short-term test-retest reliability of resting state fMRI metrics in children with and without attention-deficit/hyperactivity disorder. *Dev Cogn Neurosci*, 15, 83-93.

Song, J., Desphande, A. S., Meier, T. B., Tudorascu, D. L., Vergun, S., Nair, V. A., Biswal, B. B., Meyerand, M. E., Birn, R. M., Bellec, P., Prabhakaran, V. (2012). Age-related differences in test-retest reliability in resting-state brain functional connectivity. *PLoS One*, 7(12):e49847.

Sporns, O. (2014). Contributions and challenges for network models in cognitive neuroscience. *Nat Neurosci, 17*(5), 652-660.

Stam, C. J. (2014). Modern network science of neurological disorders. *Nat Rev Neurosci, 15*(10), 683-695.

Stephan, K. E., Penny, W. D., Moran, R. J., den Ouden, H. E., Daunizeau, J., & Friston, K. J. (2010). Ten simple rules for dynamic causal modeling. *Neuroimage*, *49*(4), 3099-3109.

Tan, P. N., Steinbach, M., & Kumar, V. (2005). Introduction to Data Mining.

Tanabe, J., Nyberg, E., Martin, L. F., Martin, J., Cordes, D., Kronberg, E., & Tregellas, J. R. (2011). Nicotine effects on default mode network during resting state. *Psychopharmacology (Berl), 216*(2), 287-295.

Taylor, B. J., (2010). The pulmonary circulation and exercise responses in the elderly. *Semin Respir Crit Care Med., 31*(5), 528-538.

Tian, L., Jiang, T., Wang, Y., Zang, Y., He, Y., Liang, M., Sui, M., Cao, Q., Hu, S., Peng, M., Zhuo, Y. (2006). Altered resting-state functional connectivity patterns of anterior cingulate cortex in adolescents with attention deficit hyperactivity disorder. *Neurosci Lett*, 400(1-2), 39-43.

Toledo, J. B., Zetterberg, H., van Harten, A. C., Glodzik, L., Martinez-Lage, P., Bocchio-Chiavetto, L., Rami, L., Hansson, O., Sperling, R., Engelborghs, S., Osorio, R. S., Vanderstichele, H., Vandijck, M., Hampel, H., Teipl, S., Moghekar, A., Albert, M., Hu, W.T., Monge Argilés, J. A., Gorostidi, A., Teunissen, C. E., De Deyn, P. P., Hyman, B. T., Molinuevo, J. L., Frisoni, G. B., Linazasoro, G., de Leon, M. J., van der Flier, W. M., Scheltens, P., Blennow, K., Shaw, L. M., Trojanowski, J. Q. (2015). Alzheimer's Disease Neuroimaging Initiative. Alzheimer's disease cerebrospinal fluid biomarker in cognitively normal subjects. *Brain*, 38 (Pt 9), 2701-15.

Tong, Y., & Frederick, B. D. (2014). Studying the Spatial Distribution of Physiological Effects on BOLD Signals Using Ultrafast fMRI. *Front Hum Neurosci, 8*, 196.

Toro, R., Fox, P. T., & Paus, T. (2008). Functional coactivation map of the human brain. *Cereb Cortex, 18*(11), 2553-2559.

Triantafyllou, C., Hoge, R. D., Krueger, G., Wiggins, C. J., Potthast, A., Wiggins, G. C., & Wald, L. L. (2005). Comparison of physiological noise at 1.5 T, 3 T and 7 T and optimization of fMRI acquisition parameters. *Neuroimage, 26*(1), 243-250.

Triantafyllou, C., Hoge, R. D., & Wald, L. L. (2006). Effect of spatial smoothing on physiological noise in high-resolution fMRI. *Neuroimage, 32*(2), 551-557.

Turner, R., Howseman, A., Rees, G. E., Josephs, O., & Friston, K. (1998). Functional magnetic resonance imaging of the human brain: data acquisition and analysis. *Exp Brain Res, 123*(1-2), 5-12.

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., Joliot, M. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage*, 15(1), 273-89.

Uddin, L. Q., Supekar, K., & Menon, V. (2010). Typical and atypical development of functional human brain networks: insights from resting-state FMRI. *Front Syst Neurosci, 4*, 21.

Van den Heuvel, M. P., & Hulshoff Pol, H. E. (2010). Exploring the brain network: a review on resting-state fMRI functional connectivity. *Eur Neuropsychopharmacol, 20*(8), 519-534.

Van Dijk, K. R., Hedden, T., Venkataraman, A., Evans, K. C., Lazar, S. W., & Buckner, R. L. (2010). Intrinsic functional connectivity as a tool for human connectomics: theory, properties, and optimization. *J Neurophysiol, 103*(1), 297-321.

Van Dijk, K. R. A., Hedden, T., Venkataraman, A., Evans, K. C., Lazar, S. W., & Buckner, R. L. (2010). Intrinsic functional connectivity as a tool for human connectomics: theory, properties, and optimization. *J Neurophysiol, 103*, 297–321.

Van Dijk, K. R. A., Sabuncu, M. R., & Buckner, R. L. (2012). The influence of head motion on intrinsic functional connectivity MRI. *Neuroimage, 59*(1), 431–438.

Van Wijk, B. C., Stam, C. J., & Daffertshofer, A. (2010). Comparing brain networks of different size and connectivity density using graph theory. *PLoS One, 5*(10), e13701.

Vannini P., Hedden, T., Becker, J. A., Sullivan, C., Putcha, D., Rentz, D., Johnson, K. A., Sperling, R. A. (2012). Age and amyloid-related alterations in default network habituation to stimulus repetition. *Neurobiology of Aging, 33*, 1237–1252.

Veer, I. M., Beckmann, C. F., Van Tol, M. J., Ferrarini, L., Milles, J., Veltman, D. J., Aleman, A., Van Buchem, M. A., Van der Wee, N. J., Rombouts, S. A. (2010). Whole brain resting-state analysis reveals decreased functional connectivity in major depression. *Front Syst Neurosci*, 4, 41.

Vilagut, G. (2014). Test-Retest Reliability. In A. C. Michalos (Ed.), *Encyclopedia of Quality of Life and Well-Being Research* (pp. 6622-6625). Dordrecht: Springer Netherlands.

Vos, S. J., Van Rossum, I. A., Verhey, F., Knol, D. L., Soininen, H., Wahlund, L. O., Hampel, H., Tsolaki, M., Minthon, L., Frisoni, G. B., Froelich, L., Nobili, F., Van der Flier, W., Blennow, K., Wolz, R., Scheltens, P., Visser, P. J. (2013). Prediction of Alzheimer disease in subjects with amnestic and nonamnestic MCI. *Neurology*, 80(12), 1124-32.

Wang, J. H., Zuo, X. N., Gohel, S., Milham, M. P., Biswal, B. B., & He, Y. (2011). Graph theoretical analysis of functional brain networks: test-retest evaluation on short- and long-term resting-state functional MRI data. *PLoS One, 6*(7), e21976.

Wang, N., Zeng, W., & Chen, L. (2012). A fast-FENICA method on resting state fMRI data. *J Neurosci Methods, 209*(1), 1-12.

Wang, X. H., Li, L., Xu, T., & Ding, Z. (2015). Investigating the Temporal Patterns within and between Intrinsic Connectivity Networks under Eyes-Open and Eyes-Closed Resting States: A Dynamical Functional Connectivity Study Based on Phase Synchronization. *PLoS One, 10*(10), e0140300.

Weissenbacher, A., Kasess, C., Gerstl, F., Lanzenberger, R., Moser, E., & Windischberger, C. (2009). Correlations and anticorrelations in resting-state functional connectivity MRI: a quantitative comparison of preprocessing strategies. *Neuroimage, 47*(4), 1408–1416.

Welvaert, M., & Rosseel, Y. (2012). How ignoring physiological noise can bias the conclusions from fMRI simulation results. *J Neurosci Methods, 211*(1), 125-132.

Weng, S. J., Wiggins, J. L., Peltier, S. J., Carrasco, M., Risi, S., Lord, C., and Monk, C. S. (2010). Alterations of resting state functional connectivity in the default network in adolescents with autism spectrum disorders. *Brain Res*. 1313, 202–214.

Westbrook, C., Kaut Roth, C., Talbot, J. (2005). MRI in Practice, third ed. *Blackwell Publishing.*

Whitlow, C. T., Casanova, R., & Maldjian, J. A. (2011). Effect of resting-state functional MR imaging duration on stability of graph theory metrics of brain network connectivity. *Radiology, 259*(2), 516-524.

Xiong, J., Parsons, L. M., Gao, J. H., & Fox, P. T. (1999). Interregional connectivity to primary motor cortex revealed using MRI resting state images. *Hum Brain Mapp, 8*(2-3), 151-156.

Yan, C., Liu, D., He, Y., Zou, Q., Zhu, C., Zuo, X. N., Long, X., Zang, Y. (2009). Spontaneous brain activity in the default mode network is sensitive to different resting-state conditions with limited cognitive load. *PLoS One*, 4(5):e5743.

Yeo, B. T., Krienen, F. M., Sepulcre, J., Sabuncu, M. R., Lashkari, D., Hollinshead, M., Roffman, J. L., Smoller, J. W., Zöllei, L., Polimeni, J. R., Fischl, B., Liu, H., Buckner, R. L. (2011). The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *J Neurophysiol*, 106(3), 1125-65.

Zeng, L. L., Wang, D., Fox, M. D., Sabuncu, M., Hu, D., Ge, M., Buckner, R. L., Liu, H. (2014). Neurobiological basis of head motion in brain imaging. *Proc Natl Acad Sci USA*, 111(16), 6058-62.

Zhang, H., Zhang, Y. J., Duan, L., Ma, S. Y., Lu, C. M., & Zhu, C. Z. (2011). Is resting-state functional connectivity revealed by functional near-infrared spectroscopy test-retest reliable?. *Journal of biomedical optics*, *16*(6), 067008-067008.

Zhou, J., Greicius, M. D., Gennatas, E. D., Growdon, M. E., Jang, J. Y., Rabinovici, G. D., Kramer, J. H., Weiner, M., Miller, B. L., Seeley, W. W. (2010). Divergent network connectivity changes in behavioural variant frontotemporal dementia and Alzheimer's disease. *Brain*, 133(Pt 5), 1352-67.

Zhou, Y., Dougherty, J. H., Jr., Hubner, K. F., Bai, B., Cannon, R. L., & Hutson, R. K. (2008). Abnormal connectivity in the posterior cingulate and hippocampus in early Alzheimer's disease and mild cognitive impairment. *Alzheimers Dement, 4*(4), 265-270.

Zuo, X. N., Kelly, C., Adelstein, J. S., Klein, D. F., Castellanos, F. X., & Milham, M. P. (2010). Reliable intrinsic connectivity networks: test-retest evaluation using ICA and dual regression approach. *Neuroimage, 49*(3), 2163-2177.

Zuo, X. N., & Xing, X. X. (2014). Test-retest reliabilities of resting-state FMRI measurements in human brain functional connectomics: a systems neuroscience perspective. *Neurosci Biobehav Rev, 45*, 100-118.

References

Zuo, X. N., Xu, T., Jiang, L., Yang, Z., Cao, X. Y., He, Y., Zang Y. F., Castellanos F. X., Milham M. P. (2013). Toward reliable characterization of functional homogeneity in the human brain: preprocessing, scan duration, imaging resolution and computational space. *Neuroimage, 65*, 374–386.