Evaluation and optimization of biologically meaningful dimensionality reduction approaches for MRI data

Inaugural dissertation

for the attainment of the title of doctor
in the Faculty of Mathematics and Natural Sciences
at the Heinrich Heine University Düsseldorf

presented by

Deepthi Varikuti
from Amalapuram

Düsseldorf, January 2018

List of Abbreviations

Acc	Accuracy
AD	Alzheimer's disease
ADNI	Alzheimer's disease neuroimaging initiative
ALE	Activation likelihood estimation
AM	Autobiographic memory
ANATICOR	Anatomy-based correlation corrections
ANOVA	Analysis of variance
aRI	adjusted rand index
AUC	Area under the Curve
bAcc	balanced accuracy
BASC	Bootstrap analysis of stable clusters
BIS	Behavioural inhibition system
BOLD	Blood oxygenation level dependent
BrainAGE	Brain age gap estimation
CBMA	Coordinate-based meta-analysis
CogAC	Cognitive action control
CSF	Cerebrospinal fluid
CV	Cross Validation
DA	Dopamine
DemTect	Early dementia detection assessment
DMN	Default mode network
DVARS	D referring to temporal derivative of timecourses, VARS referring to RMS variance over voxels
DLPFC	Dorsolateral prefrontal cortex
EC	Eyes closed
EmoSF	Emotional scene and face processing network
Emp	Empathy
EO	Eyes open
EPI	Echo-planar Imaging
eSAD	Extended socio-affective default mode network

Face	Static face perception
FC	Functional connectivity
FD	Frame-wise displacements
FFFS	Fight-Flight-Freeing System
FFM	Five Factor Model
FIX	FMRIB's ICA-based X-noisifier
fMRI	Functional magnetic resonance imaging
FWHM	Full width at half maximum
GM	Grey matter
GMV	Grey matter volume
GrpGM	Group grey matter mask
GSR	Global signal regression
HC	Healthy Control
HCP	Human connectome project
HCSCZ	Healthy Control from Schizophrenia dataset
HCPD	Healthy Control from Parkinson's dataset
H & Y Scale	Hoehn and Yahr Scale
IC	Independent component
ICA	Independent component analysis
ICA-AROMA	Independent component analysis based Automatic Removal Of Motion Artifacts
LaIns	Left anterior insula
ICC	Intra class correlations
ICD-10	International Classification of Disease
IndGM	Individual grey matter mask
IOG	Inferior occipital gyrus
LASSO	Least absolute shrinkage and selection operator
LOO-CV	Leave-one-out approach
MA	Modeled activation map
MAE	Mean absolute error
MCI	Mild cognitive impairment

MELODIC	Multivariate exploratory linear optimized decomposition into independent components
MNI	Montreal neurological institute
MNS	Mirror neuron system
MoCA	Montreal cognitive assessment
Motor	Motor execution
MRI	Magnetic resonance imaging
MVPA	Multivariate pattern analysis
NA	Normal Aging
NAc	Nucleus accumbens
NEO-FFI	NEO Five-Factor Inventory
NINCDS/ADRDA	National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer's Disease and Related Disorders Association
NNMF	Nonnegative matrix factorization
NNSVD	Non-negative double singular decomposition
NoGM	No grey matter mask
NoPCA	No PCA denoising
OFC	Orbitofrontal cortex
OPNMF	Orthonormal projective non-negative matrix factorization
Pain	physiological stress/pain
PANSS	Positive and Negative Symptom Scale
PCA	Principal component analysis
PD	Parkinson's disease
PET	Positron emission tomography
PFC	Prefrontal cortex
RaIns	Right anterior insula
Rew	Reward
ROC	Receiver operating characteristics

RoCO	Reliability at the connection level
ROI	Regions of interest
RoSO	Reliability at the subject level
RS	Resting-state
RSFC	Resting-state functional connectivity
RST	Gray's reinforcement sensitivity theory
RVM	Relevance vector machine
SCZ	schizophrenia
SCID-P	structured clinical interview
SD	Standard deviation
Sens.	Sensitivity
SM	Semantic memory
sMRI	Structural magnetic resonance imaging
Spec.	Specificity
SPL	Superior parietal lobule
SPM	Statistical parametric mapping
SPSS	Statistical package for the social sciences
SVD	Singular value decomposition
SVM	Support vector machine
ToM	Theory-of-mind cognition
TPQ	Cloninger's Tridimensional Personality Questionnaire
TSR	Tissue signal regression
UPDRS	Unified Parkinson's Disease Rating Scale
VA	Vigilant attention
VBM	Voxel based morphometric
VOI	Volume of interest
WM	White matter
WMCSF	White matter and Cerebral spinal fluid signal regression
WMN or WM network	Working-memory network

TABLE OF CONTENTS

Zusammenfassung

Multivariate Merkmalsanalysemethoden wurden bisher sehr häufig in den Neurowissenschaften eingesetzt, um eine Beziehung zwischen unterschiedlichen phänotypischen und behavioralen Messgrößen und bildgebenden neuronalen Korrelaten herzustellen. In diesem Zusammenhang ist man zusätzlich mit dem Problem konfrontiert, dass hochdimensionale voxelbasierte bildgebende Daten eine hohe Rechenleistung erfordern und gleichzeitig ein schwaches Signal-Rausch Verhältnis aufweisen. Des Weiteren bieten MRT basierte bildgebende Verfahren die Möglichkeit, funktionale und anatomische Informationen des Gehirns, die auf unterschiedliche Modalitäten beruhen, zu erfassen. Daher war das Ziel dieser Arbeit, den Fluch der Dimensionalität in Verbindung mit hochdimensionalen Voxel-Level-Informationen durch die Optimierung und Bewertung verschiedener Methoden auf der Basis von funktionellen und strukturellen MRT-Daten zu untersuchen. In diesem Rahmen wurden besonders die Aspekte der Reliabilität und Interpretierbarkeit betrachtet. Der erste Teil des Projekts konzentrierte sich auf die Untersuchung der Effekte von verschiedenen Verfahren zur Beseitigung von konfundierten Variablen auf die Test-Retest Reliabilität der funktionalen Konnektivitätsschätzungen von a priori definierten funktionalen Netzwerken. Im Allgemeinen liefert eine voxelbasierte Berechnung der funktionalen Konnektivität des gesamten Gehirns eine Vielzahl von Konnektivitätsmaßen (z.B. 300000 x 300000). Dabei sind jedoch für eine gegebene interessierende Variable nur wenige dieser Verbindungen in der multivariaten Merkmalsanalyse informativ. Aufgrund dessen sind a priori definierte seed-basierte FC Messungen rechnerisch weniger aufwendig und informativer. Die beste Strategie zum Extrahieren der Konnektivitätsmatrizen aus a priori Netzwerken bleibt jedoch immer noch ungeklärt. Daher konzentriert sich Studie 1 hauptsächlich auf das Erreichen zuverlässiger funktionaler Konnektivitätsmaße. Weiterhin untersuchten wir unsere Annahmen in Studie 2 und 3 mit funktionellen MRT-Daten, indem wir die Merkmalsreduktion auf Grundlage von Fachwissen (d.h. a priori definierte meta-analytische Netzwerke) und unter der Anwendung der multivariaten Merkmalsanalyse implementierten. Der letzte Teil des Projekts zielte darauf ab, eine alternative Dimensionsreduktionmethode zu implementieren, nämlich die nicht-negative Matrixfaktorisierung (NNMF). Im Gegensatz zu der häufig verwendeten Hauptkomponentenanalysemethode erhöht die neue Methode die Qualität der Interpretierbarkeit der Niedrigrang-Approximation. Hauptsächlich wurde NNMF auf die voxelbasierte morphometrische (VBM) Messung angewendet, die aus anatomischen MRT Daten berechnet wurde. In diesem zweiten Teil des Projekts wurde die NNMF-basierte Reduktion von VBM-Daten zur Vorhersage des Alters evaluiert. Darüber hinaus ermittelten wir die Eigenschaften der Gehirnregionen, die in die Vorhersageanalyse eingeflossen sind.

Zusammenfassend wurde in Studie 1 überprüft, welchen Einfluss verschiedene Confound-Entfernungsverfahren und Signalextraktionsansätze auf die Reliabilität von funktionalen Konnektivitätswerten in a priori definierten kanonischen Netzwerken haben. Die drei folgenden Arbeiten bewerteten verschiedene Merkmalsreduktionsstrategien in auf Machine Learning basierenden Studien, die entweder funktionelle oder strukturelle MRT-Daten verwendeten. Die Ergebnisse dieser Studien veranschaulichen effiziente und zuverlässige Strategien, wie der hochdimensionale Merkmalsraum in biologisch plausibel abgegrenzte Merkmale für die Analyse von Gehirn-Verhalten oder Gehirn-Phänotyp reduziert werden kann.

Abstract

Multivariate approaches have been increasingly applied in the field of neuroscience, to relate neuroimaging pattern to various phenotypical or behavioral measures. In this context, high-dimensional voxel wise neuroimaging data lead to multiple issues such as heavy computational demand, as well as poor signal to noise ratio. In addition, (Magnetic resonance imaging) MRI based neuroimaging offers the opportunity of capturing the functional and anatomical information of the brain based on different modalities. Therefore, in this project, we aimed to address the curse of dimensionality associated with high-dimensional voxel level information by optimizing and evaluating various methods used within either functional or structural MRI data, giving particular attention to reliability and interpretability. The first part of the project focused on investigating the effects of various confounds removal approaches on the test-retest reliability of functional connectivity estimates of a-prior defined functional networks. In general, voxel-wise whole brain functional connectivity computation provides a multitude of connectivity measures (i.e., 300000 x 300000), while only few of these connections are informative during the multivariate approaches for a given variable of interest. In turn, a priori defined seed based FC measures are computationally more tractable and more informative. However, the best strategy to extract the connectivity matrices from a priori networks still remains as a question. Hence, study 1 mainly focuses on achieving reliable functional connectivity measures. Study 2 and study 3 then further examined our assumptions by implementing the feature reduction based on domain knowledge (i.e., a priori defined meta-analytic networks) during the investigation of the multivariate approaches implicated to the given target function using the functional MRI data. The last part of the project aimed to implement an alternative dimensionality reduction method namely, non-negative matrix factorization, which promotes the quality of interpretability of the low-rank approximations in contrast to the frequently used principal component analysis method. Mainly, Non-negative matrix factorization (NNMF) was applied to voxel based morphometric (VBM) measure computed from anatomical MRI data. In this second part of the project, NNMF based reduction of VBM data has been evaluated for prediction of age. Furthermore, we investigated the patterns of brain regions contributed in the prediction analysis. In sum, study 1 has been dedicated to investigate the influence of various confound removal procedures and signal extraction approaches on the reliability of the functional connectivity scores in a priori defined canonical networks. The three following studies assessed various feature reduction strategies in machine learning based studies using either functional or structural MRI data. These results of these studies demonstrated efficient and reliable strategies to reduce the high dimensional feature space into biologically plausible confined features for brain-behavior/phenotype relationship analyses.

General Introduction

1 Basics of the concepts and techniques implemented in the study

Here in this section, the fundamentals of the concepts and techniques used within the thesis are explained. However, a detailed explanation of the rationale of the study is presented in the next section. In this section, first the basics of Magnetic resonance imaging are introduced. Later a brief introduction of data reduction and multivariate approaches implemented in this thesis are presented.

**1.1 The Physics of Magnetic Resonance Imaging**

Magnetic resonance imaging (MRI) relies on the signal acquired from the nuclear resonance properties of hydrogen atoms present in the water molecules of the tissues (Brown et al. 2004; Weishaupt et al. 2007). In the presence of an external magnetic field, hydrogen atoms are aligned either in the direction or opposite direction to the magnetic field. In addition, nucleus present in an atom consists of a property known as precession, which means the atom is constantly spinning about an axis at a constant rate. The frequency of precession is proportional to the strength of the magnetic field as expressed in the Larmor equation.

$$w_0 = \gamma \, {B_0}/{2\pi}$$

where w_0 is the Larmor frequency, B_0 is the magnetic field strength and γ is a constant, known as gyromagnetic ratio. In order to activate the protons into higher energy state, a radio frequency pulse is applied in the perpendicular direction to the magnetic field. Thus, the hydrogen atom experiences a flip away from direction of the magnetic field. Proton in the excited state returns into the equilibrium state by emitting the excess energy into the surrounding. MRI is based on this energy generated by the proton, known as MR signal (Brown et al. 2004). The MR signal can be measured based on processes known as spin-lattice interaction and spin-spin interaction triggering T1 relaxation and T2 relaxation, respectively.
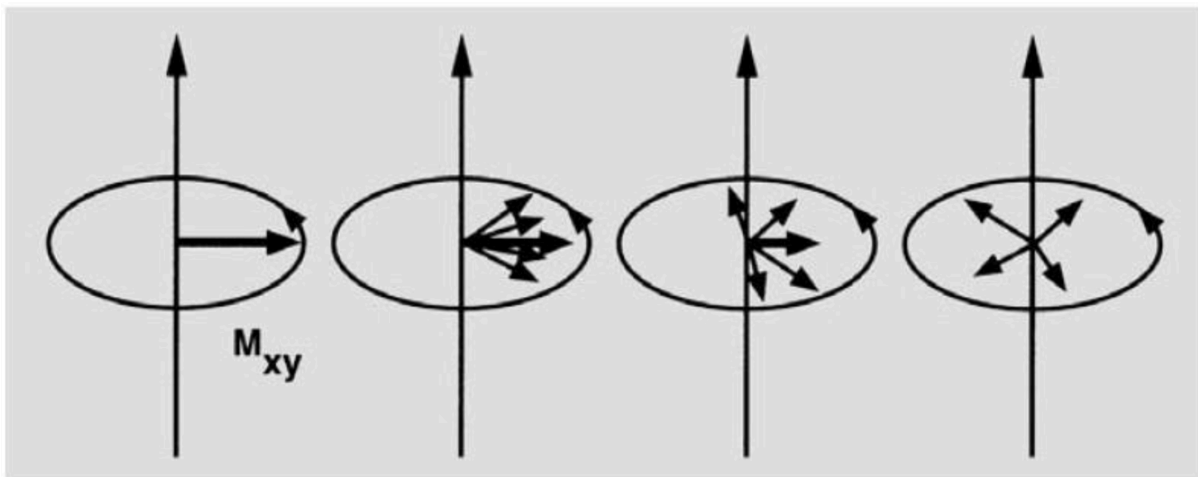
T1: Longitudinal Relaxation:

After switching off the radio frequency, the proton flipped away from the magnetic field returns into the normal condition. The time constant for the recovery of the longitudinal component (i.e., longitudinal relaxation) is called T1. This process is called as spin-lattice interaction as the excess energy is released into the surroundings.
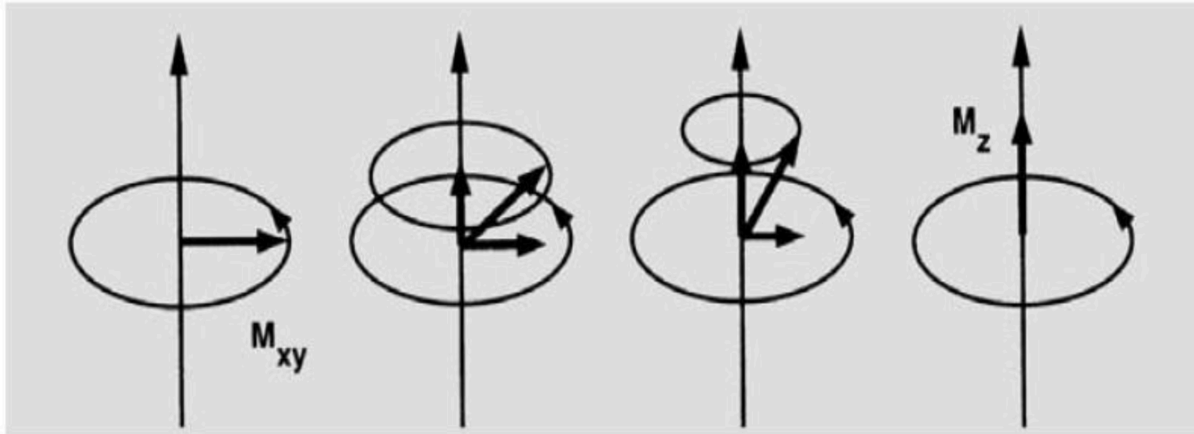
T1 relaxation. Re-growth of longitudinal magnetization requires an exchange of energy. M_{xy} is the longitudinal component of the magnetization and M_z is the transversal component of the magnetization (Weishaupt et al. 2007).

T2/T2*: Transverse Relaxation:

Initially, all the neighboring atoms exhibit coherent precession (i.e., to spin in phase with others) in the direction of the external magnetic field. However, immediately after excitation, the atoms experience a flip into the transversal plane (i.e., perpendicular to the external field direction), however, preserve the phase coherence. The interaction that all atoms eventually have with each other, generates a loss of the phase coherence due to the local variation in their magnetic field environment. Hence, magnetization in the transverse plane is lost. The time constant for the decay of this transverse component (i.e., transversal relaxation) is called T2. Besides, it is also possible to define a time constant for the decay of transverse component occurred due to the inhomogeneities induced by the local magnetic field, called T2*. This process is called spin-spin interaction as the exceeding energy from one proton is transferred to the neighboring proton.

T2 and T2* relaxation. Lose phase coherence resulting in the loss of transverse magnetization without energy dissipation. M_{xy} is the longitudinal component of the magnetization (Weishaupt et al. 2007).

1.2 Signal contrast and voxel-wise measurement in structural and functional modalities of MRI

MRI based neuroimaging offers the opportunity of capturing the anatomical and functional information of the brain based on different modalities. These modalities allow us to examine the most fundamental question of the neuroscience community. The former modality answers how does the spatial representation of the brain characterize the structural information, the latter allows us to understand how does brain function and further understand how this information alters over time.

First, Structural magnetic resonance imaging (sMRI) is a technique, which identifies differences among tissues of the brain and further measures the density of each tissue (Symms et al. 2004). Variation in the water content among tissue types originates differences in the relaxation timings. Thereby, different relaxation rates produce substantial signal contrast among tissues to achieve significant morphological representation (Brown et al. 2004; Weishaupt et al. 2007). For instance, cerebrospinal fluid with more water content has a longer T2 compared to grey matter. Thus, differentiation of various tissue types based on T1 or T2 weighted imaging is rather evident (for example, on a T1-weighted image the cortex appear brighter than ventricles) to measure the structural changes of the brain. Usually, sMRI is used to explore the anatomy of the brain, measured based on T1-weighted imaging. Here, each voxel of the brain image contains the volume of grey matter at that spatial location. Voxel based morphometry (VBM) is one of the most commonly used methods to measure grey matter volume (Good et al. 2001). It provides structural measures, which convey biologically meaningful information and capture brain changes related to age and pathology, as well as brain plasticity (Good et al. 2001; Tisserand et al. 2002; May 2011). Structural MRI data was preprocessed with the VBM8 toolbox (http://www.neuro.uni-jena.de/vbm8) to derive voxel-wise grey matter volumes for each subject. A detailed explanation of this procedure is explained in the section 2.2 of study 4.

Furthermore, Functional magnetic resonance imaging (fMRI) is a technique of brain imaging, which relies on identification of the blood oxygenation fluctuations associated with the neural activity (Buxton 2002; Scott A. Huettel; Allen W. Song; Gregory McCarthy et al. 2004). fMRI is widely used to study alterations in brain functions between healthy and clinical disordered subjects In the absence of oxygen,

the field inhomogeneities are enhanced due to the paramagnetic property of the hemoglobin (Ogawa et al. 1992). Therefore, deoxygenated hemoglobin exhibits faster decay of $T2^*$ due to the increased de-phasing of the water molecules. When a particular region is functionally active, brain tissue in that region encounters an increase in the flow of oxygenated blood and further leads to a longer $T2^*$. Thus, fMRI measures the brain activity based on the concentration of the oxygenated blood (also known as Blood Oxygenation Level Dependent (BOLD) signal), which is sensitive to $T2^*$ and further investigates how this activity (i.e., BOLD signal) fluctuates over time (Schölvinck et al. 2010). To note, several confounding factors (such as system noise, thermal noise and noise induced by non-neuronal physiological processes) may impact the BOLD signal and hence apparently brain activity. Therefore, fMRI data was preprocessed using SPM8, (www.fil.ion.ucl.ac.uk/spm) to derive voxel-wise BOLD signal over time for each subject. The functional connectivity between two regions (or voxels) was then computed as the correlation coefficient between these time series, which were transformed to Fisher's Z scores to render them normally distributed. A detailed explanation of this procedure is explained in the section 2.2.2 of study 1.

1.3 Fundamentals of various techniques used in this thesis

Univariate analysis has been the classical approach applied at voxel level, to investigate the group differences in the brain (Bandettini et al. 1992; Friston et al. 1994; Friston 1997; Mukamel et al. 2005). Despite the approach proved to effectively provide inferences about the brain regions associated with a given task, the results are merely specific to the type of research question examined (i.e., univariate analysis is generally used for hypothesis-driven study while multivariate analysis is generally more exploratory). In addition, univariate voxel-wise analysis infers on identical information from several single voxels (Gonsalves and Cohen 2010) and thus fails to explicate patterns based on integrated information from multiple voxels, even when the voxels share non-identical variance (i.e., univariate analysis fails to identify complex relationships between different regions). In contrast, multivariate analyses aim to estimate a given phenotypical or behavioral measure by capitalizing on the spatially distributed patterns over a set of voxels (Haxby et al. 2001; Kamitani and Tong 2005; Pereira et al. 2009). Given the advantage of multivariate analysis exploiting information from several voxels with different properties, the results obtained with these approaches remain more stable and can be analyzed in several directions (O'Toole et al. 2007; Yoon et al. 2008; Habeck 2010). Thus, multivariate approaches have been increasingly applied in the field of neuroscience, to investigate neuroimaging patterns associated with various phenotypical or behavioral measures (Cox and Savoy 2003; Craddock et al. 2009; Franke et al. 2010; Davatzikos et al.

2011). However, the efficiency of the multivariate approaches greatly relies on multiple dimensions of the neuroimaging data. Briefly, there are three dimensions of the neuroimaging data: number of variables (number of voxels), number of observations (sample size), and multitude of modalities (data from different modalities). The current study attempted to address the challenges associated with these dimensions, in particular to provide solutions to meaningfully reduce the dimensionality associated with the number of voxels, in the context of multivariate approaches. Thus, a brief introduction is presented below about the different data reduction techniques with different multivariate approaches used in this dissertation.

1.3.1 Dimensionality reduction techniques

1.3.1.1 Coordinate-based meta-analysis

Coordinate-based meta-analysis (CBMA) yields a quantitative summary of brain regions identified to answer a certain research question, exploiting the reported results (such as, coordinates of the activated regions) across neuroimaging literature. Therefore, networks derived using a coordinate-based meta-analysis technique have been used as the ground truth to investigate the reliability of the connectivity measure using different preprocessing methods in study 1. Furthermore, coordinate-based meta-analysis technique has been mainly used to reduce the features based on existing 'domain knowledge' in study 2 and study 3.

In order to obtain CBMA networks, revised activation likelihood estimation (ALE) algorithm is performed (Turkeltaub et al. 2002; Laird et al. 2005; Eickhoff et al. 2009; Yarkoni et al. 2011; Eickhoff et al. 2012) employing MATLAB packages. Normally, ALE aims to capture the uncertainty in spatial location related to each of the reported coordinate, by modeling an isotropic Gaussian probability distribution around the reported foci. Moreover, the size of the Gaussian kernel is modeled by accounting the spatial uncertainty (estimated by computing the Euclidean distance) of both between-template variance (i.e., bias initiated by different normalization strategies) and between-subject variance (i.e., smaller sample sizes based bias). These empirical estimates can be described by the Maxwell-Boltzmann distribution, which exhibits isotropic normal distribution. The underlying isotropic normal distribution (i.e., displacement in X, Y, Z directions) has zero-mean and standard deviation (denoted as α) based on the Maxwell-Boltzmann distribution, where α-parameter resembles the standard deviation of the underlying Gaussians displacement (i.e., desired kernel size, denoted as σ). The point-estimate (μ) of Maxwell-Boltzmann distribution is replaced by the Euclidean distance computed from our data, and derived the α-parameter (hence the σ of the Gaussian displacement) by solving the following equation.

$$\sigma_{\text{sub}} = \alpha_{\text{sub}} = \frac{\overline{ED_{\text{sub}}}}{2 * \sqrt{\frac{2}{\pi}}}$$

$$\sigma_{\text{temp}} = \alpha_{\text{temp}} = \frac{\overline{ED_{\text{temp}}}}{2 * \sqrt{\frac{2}{\pi}}}$$

Here, ED_{sub} is the mean Euclidean Distance between corresponding foci of different subjects and ED_{temp} is the mean Euclidean Distance between corresponding maxima as observed in the different group-analyses (due to the different normalization strategies).

The corresponding Full width at half maximum (FWHM)) parameter is assessed as following, using the σ of a Gaussian distribution.

$$FWHM_{sub} = \sigma_{sub} * \sqrt{8 * \log(2)}$$

$$FWHM_{temp} = \sigma_{temp} * \sqrt{8 * \log(2)}$$

Influence of the inter-subject variability on the spatial uncertainty in a group of N subjects can be measured as

$$FWHM_{sub\ (effective)} = \frac{FWHM_{sub}}{\sqrt{N_{subjects}}}$$

Subsequently, the uncertainty of the spatial location of a given coordinate is modeled by combining the empirical estimations of the between-template and between-subjects variance outlined above.

$$FWHM_{effective} = \sqrt{(FWHM_{temp})^2 + \left(\frac{FWHM_{sub}}{\sqrt{N_{subjects}}}\right)^2}$$

Thus, for a given study, modeled activation map is computed by integrating the probabilities associated to all the foci of that particular study (specific to one) (i.e., $1 - \prod_i^{N foci}(1 - p(i))$, Here, $p(i)$ is the probability of ith focus at a given voxel. Finally, combining these MA maps across studies reveals convergence of results from

multiple studies at a particular brain location. Thus, the most probable regions associated to the given research question is achieved.

1.3.1.2. Principal component analysis

Principal component analysis (PCA) is one of the most commonly used methods with multiple applications. For instance, in this dissertation, PCA is used as de-noising strategy in study 1, and as dimensionality reduction technique in study 4. Typically, PCA aims to achieve reduced representation with uncorrelated features explaining highest variation in the dataset, which are linear combinations of the original correlated features (Jolliffe 2005; Mourão-Miranda et al. 2005; Mourão-Miranda et al. 2012). PCA thus decomposes the entire original representation into low rank approximations with a combination of positive and negative weights, which does not promote spatially localized components. PCA decomposition captures the component with highest variance as the most dominant principal component, followed by maximum fit of the remaining variance by the successive components (Jolliffe 2002).

Principal components (W_x) are a linear combination of the original parameters:

$$W_x = \sum_{i=1}^{N_{tot}} a_{ix} V_i$$

Where $-1 < a_{ix} < 1$ are the coefficients of the linear transformation, V_i are the original features and N_{tot} is the number of original features. In other words, PCA minimizes the following optimization problem:

$$\min_{W,H} \parallel V - WH \parallel_F^2$$
$$subject\ to\ H = W^T V$$
$$W^T W = I$$

Here, W indicates the component space and H represents the loading coefficients of the components. The original features (V) are decomposed using singular value decomposition (SVD) into eigenvalues of the covariance matrix. Capitalizing on the sample size of the data, most relevant components are computed and denoted as

principal components. However, the signed components within the PCA decomposition engage complex cancellations during the reconstruction of the original representation. Therefore, the use of PCA-based dimensionality reduction on brain voxels results in hardly interpretable components, which can in turn prevent the interpretation of a predictive model based on PCA-derived components.

1.3.1.3. Non-negative matrix factorization

Non-negative matrix factorization (NNMF) has recently been suggested as a plausible factorization of high-dimensional VBM data. Non-negativity and sparsity the components trend to reflect distinct, anatomically interpretable "building blocks" rather than whole-brain patterns of positive and negative values obtained from, e.g., PCA, (Lee and Seung 1999). Therefore, in study 4, we use the same orthonormal projective non-negative matrix factorization (OPNMF) approach as described by (Sotiras et al. 2015) that reduces computation time and yields deterministic solutions (Yang et al. 2007; Yang and Oja 2010a) OPNMF factorized the data 'X' into two non-negative sub matrices (W and H) representing the sparse components (the dictionary) and the subject-specific loading coefficients in the ensuing low-rank space by solving the following optimization problem through minimization of the squared Frobenius norm (i.e., reducing the reconstruction error).

$$\min_{W \geq 0, H \geq 0} \parallel X - WH \parallel_F^2$$
$$subject\ to\ H = W^T X$$
$$W^T W = I$$

As stated, OPNMF is an extended framework of the standard NNMF. The constraint $H = W^T X$, was one of the extended constraint (i.e., the projective basis function). The loading coefficient matrix was estimated by projecting the input matrix onto the estimated component matrix. Estimating H using projection basis function would provide more localized representation, as well as decline the overlap between the estimated components (i.e., reduce the risk of overcomplete low rank representation without exceeding the dimensionality of the input, which allows us to accurately reconstruct the original representation (Yang et al., 2007)). In addition, this constraint has improved the representation by inducing the sparsity into the component matrix 'W' (Yang et al., 2007, Yang and Oja, 2010). In addition, $W^T W = I$ promotes the orthogonality which allows the components to be independent of each other. Yang and Oja 2010 mathematically demonstrated that the addition of orthonormal constraint reduced computational complexity.

In order to find the solution, NNMF employed multiplicative update rule, which facilitates the non-negativity into the solutions. However, the multiplicative update rule has been modified as reported by Yang and Oja 2010, in order to satisfy the additional constraints of orthonormal projection basis function. The following update rule has been iteratively applied until it converges to achieve an optimal solution.

$$W'_{ij} = W_{ij} \frac{(XX^TW)_{ij}}{(WW^TXX^TW)_{ij}}$$

Where, i = 1…Number of voxels, j = 1…Number of components. Prior to the optimization scheme, a dual initialization step termed as non-negative double singular decomposition (NNSVD) has been implemented (Boutsidis and Gallopoulos 2008). This initialization step aimed to provide a good approximation by reducing the residual error relatively faster, as well as to enable reproducible final solutions. Furthermore, initialized sparse components ensure the final sparsity. Lastly, NNSVD initialization attempts to enable reproducible solutions across several runs.

To summarize the factorization process, W is first initialized through non-negative double singular value decomposition. Later, W is iteratively updated with the multiplicative update rule, until it converges to an optimal solution. Finally, projecting X onto W to obtain a solution that minimizes the reconstruction error yields H.

1.3.2 Multivariate approaches for prediction or classification

1.3.2.1 Support Vector Machine

Support Vector Machines (SVM) are a popular machine learning method for classification (Vapnik 1998). In study 2, we implemented classification-based machine learning approach to classify patients and healthy controls into separate groups. Hence, we implemented SVM in this study 2. In study 2, non-sparse linear two-class SVMs were computed using LibSVM (Chang and Lin 2011) [Chang and Lin, 2011] (https://www.csie.ntu.edu.tw/~cjlin/ libsvm). When the training dataset of n points are given in the form, $(x_1, y_1) \dots (x_n, y_n)$, where $y_i \epsilon \{1, -1\}$, indicating the class each point x_i belongs. SVM solves the following optimization problem, which aims to find a hyper plane that divides the two groups.

$$\min_{w,b,\xi} \frac{1}{2}w^Tw + C\sum_{i=1}^{n} \xi_i$$
$$subject\ to \quad y_i(w^T\Phi(x_i) + b) \geq 1 - \xi_i$$

$$\xi_i \geq 0, i = 1, \dots n$$

Where w is the normal vector to the hyperplane; $\Phi(x_i)$ maps x_i into a higher dimensional space and $C > 0$ gives the regularization factor. Usually, the following dual optimization problem is solved, given the high dimensionality of the hyperplane vector (w).

$$\min_{\alpha} \frac{1}{2}\alpha^T Q\alpha - e^T\alpha$$
$$subject\ to\quad y^T\alpha = 0$$
$$0 \leq \alpha_i \leq C, i = 1, \dots n$$

where e $= [1, \dots, 1]^T$ is the vector of all ones, Q is a positive semidefinite matrix with entries $Q_{ij} = y_i y_j K(x_i, x_j)$, and the kernal functions are $K(x_i, x_j) \equiv \Phi(x_i)^T \Phi(x_j)$. After solving the aforementioned dual optimization problem, using the primal-dual relationship, the optimal w satisfies

$$w = \sum_{i=1}^{n} y_i \alpha_i \Phi(x_i)$$

Finally the decision function is

$$sign(w^T\Phi(x_j) + b) = sign\left(\sum_{i=1}^{n} y_i\alpha_i K(x_i, x) + b\right)$$

In the model for prediction or classification, the following parameters are saved, $y_i, \alpha_i, K(x_i, x)$ $(i.e., kernel\ parameters), b$, support vectors and label names.

1.3.2.2 Relevance Vector Machine

The Relevance Vector Machine (RVM) is another machine learning approach, which was developed from support vector machine, but provides solutions employing probabilistic Bayesian learning (Tipping 2001; Tipping and Faul 2003a). Most importantly, in contrast to SVM, RVM induces sparsity into the regression model, which ultimately overcomes the over fitting issue for high dimensional data. Therefore, RVM is implemented as the sparse regression model to predict various personality traits in Study 3. For doing so, statistical learning of the sparse regression model employing RVM was implemented using the SparseBayes package (http://www.miketipping.com/index.htm). When the training dataset of n points are

considered in the form$(x_1, y_1) \ldots (x_n, y_n)$. The following generalized linear model is trained on the training dataset.

$$f(x) = \sum_{i=1}^{n} w_i k(x, x_i) + \varepsilon$$

where $x = [x_i, \ldots, x_n]^T$, $k(x, x_i)$ is a bivariate kernel function centered on each of the training data points, $w = [w_i, \ldots, w_n]^T$ are the regression coefficients and ε is the noise. Here, the output is assumed to follow a Gaussian distribution with mean y(x$_i$) and uniform variance σ^2 of the noise ε, so $p(t|x) = N(t|y(x), \sigma^2)$. Conditional probability of the target variables following those assumptions is expressed as

$$p(t|w, \sigma^2) = (2\pi\sigma^2)^{-N/2} \, exp\left\{-\frac{1}{2\sigma^2}\|t - \Phi w\|^2\right\}$$

Here, the kernel function matrix Φ represents all the pairs $\Phi_{i,j} = K(x_i, x_j), i, j \, \epsilon \, [1, \ldots, n]$. In order to induce the sparsity, an additional vector of hyper parameters (α_i) is introduced to parametrize the width of the normal prior distribution.

$$p(w_i|\alpha_i) = \sqrt{\frac{\alpha_i}{2\pi}} \, exp\left(-\frac{1}{2}\alpha_i w_i^2\right)$$

The goal of the RVM is to iteratively solve (the following steps) the type II maximization of the marginal likelihood $p(t|\alpha, \sigma^2)$ with respect to α and σ^2, which reduces the dimensionality of the problems when any of the α_i is larger than the defined threshold. Finally, the algorithm stops, when there is no further improvement in the likelihood $p(t|\alpha, \sigma^2)$ (Tipping 2001).

$$\alpha_i^{new} = \frac{1 - \alpha_i \Sigma_{ii}}{\mu_i^2}$$
$$(\sigma^2)^{new} = \frac{\|t - \phi\mu\|^2}{n - \sum_{i=1}^{n}(1 - \alpha_i \Sigma_{ii})}$$

The unknowns are computed as following, where, $B = \sigma^{-2} I_{NxN}$ and $A = A = diag(\alpha_0, \alpha_1 \ldots \alpha_N)$.

$$\Sigma = (\phi^T B \phi + A)^{-1}$$
$$\mu = \Sigma\phi^T B t$$

1.3.2.3 Regularized (sparse) regression model

As an alternative approach to RVM, LASSO (least absolute shrinkage and selection operator) has been the most commonly implemented sparse regression model in the machine learning studies. LASSO is known to perform both variable selection and regularization to improve the precision and interpretability of the prediction model (Tibshirani 1996). Regularization, in general, is to introduce an additional penalty term, which can improve the generalizability of the learning model. Therefore, in study 4, we performed an additional comparison between LASSO and RVM regression models, given that both the methods induce sparsity into the learning model. However, in study 4, LASSO for learning a (sparse) linear regression model predicting the subjects' age is implemented as in the 'glmnet' package, https://www.jstatsoft.org/article/view/v033i01 (Tibshirani 1996). To note, elastic net regularization based framework is employed within the 'glmnet' package, which can easily be adapt into purely LASSO regularized framework. The mathematical definition of Elastic net is explained below. Elastic net regularization is alternative regularization and variable selection model with sparsity of representation similar to LASSO (Zou and Hastie 2005).

When the training dataset of n points are considered in the form $(x_1, y_1), \dots, (x_n, y_n)$ the response y is predicted as following using a regular linear regression model.

$$\hat{y} = \widehat{\beta_0} + x_1 \widehat{\beta_1} + \cdots + x_k \widehat{\beta_k}$$

In the model fitting procedure using ordinary least square (OLS), a vector of coefficients ($\hat{\beta} = (\widehat{\beta_0} \dots \widehat{\beta_k})$) is estimated by minimizing the residual sum of squares (where, k is the number of features). However, to improve the prediction accuracy and interpretation of the learning model, penalized least square techniques have been introduced (Tibshirani 1996; Zou and Hastie 2005), which introduces a regularization term to the ordinary least square estimations (such as L1 and L2 norm regularization (Horn, R. A. and Johnson 1973)). The elastic net regularization framework that linearly combines the L1 and L2 norm regularizations, calculated the vector of coefficients by minimizing the following penalized least square function (Zou and Hastie 2005; Bunea et al. 2011).

$$\min_{\beta_0, \beta} \|y - \beta_0 - x \cdot \beta\|^2 + \lambda \left[\frac{1 - \alpha}{2} \|\beta\|_2^2 + \alpha \|\beta\|_1 \right]$$

$$subject\ to\ \|\beta\| \leq t$$

Where,

$$\|\beta\|_1 = L1 \; penalty \; term = \sum_{j=1}^{k} |\beta_j|$$

$$\|\beta\|_2^2 = L2 \; penalty \; term = \sum_{j=1}^{k} \beta_j^2$$

Here, $t \geq 0$ is a tuning parameter, which controls the amount of shrinkage applied to the estimates (i.e., some coefficients be equal to zero) and $\lambda \geq 0$ is the hyper parameter defining the amount of penalty, which controls the complexity of the regression model (i.e., balance between sparsity and high prediction accuracy). The elastic net parameter is $0 \leq \alpha \leq 1$, further defined as LASSO when $\alpha = 1$ and ridge when $\alpha = 0$. To note, sparsity of the solution is encouraged by L1 penalty term and stability of the solution is promoted by L2 penalty term (Hastie et al. 2009; Friedman et al. 2010).

2 Rationale of the study

Until recently, most of the multivariate and also univariate analysis were applied on smaller sample sizes, which have been shown to be prone to false positives (i.e., reduced chance of the effect being biologically plausible) (Button et al. 2013; Button 2014). Analysis performed on smaller sample size allows only a confined set of outcomes, which are specific to the given sample. To note, any minor methodological manipulation could easily deviate the results in such underpowered sample-sized analysis. In addition, taking the highly variant inter-individual difference into account, systematic examination of the pattern within brain regions can only be achieved by increasing the sample size. Thus, one needs to implement the multivariate approaches on very large sample size i.e., for thousands of subjects, to enhance the stability of the outcomes. Even though, recent studies began to investigate the pattern of the brain regions on relatively larger sample size datasets, yet, the sample size hasn't been more than 6000 subjects (Miller et al. 2016a). In contrast, the voxel-wise data contains more than 300000 voxels. Therefore, even with the recent experimental settings, the dimensionality associated with voxel-wise data has greatly exceeded the sample size (Guyon and Elisseeff 2003; Schrouff et al. 2013; Mwangi et al. 2014). Such high-dimensional voxel wise neuroimaging data lead to multiple issues such as heavy computational demand, as well as poor signal to noise ratio. Normally, voxel wise data is a mixture of noise, redundant, and information of interest with multivariate properties. Therefore, voxel-wise representation could potentially fit several multivariate models for the same data, (i.e. voxel-wise representation goes with 'risk of overfitting' (Guyon and Elisseeff 2003; Hua et al. 2009)). To avoid the risk of overfitting and improve the signal to noise ratio, a dimensionality reduction technique is crucially needed. Given the aim of dimensionality reduction technique is to eliminate the influence of the noisy and redundant features, the pattern of brain regions identified after data reduction are expected to offer a deeper interpretation of the research question. Therefore, the dimensionality reduction has to transform the high-dimensional data into low rank approximations, while still retaining the most influential structure of the original data (Guyon and Elisseeff 2003; Fan et al. 2007; Mwangi et al. 2014). Hence, the inevitability of data reduction associated with high dimensional voxel wise data motivated our study to evaluate different methods addressing this particular issue (i.e., dimensionality related to the number of variables The major objectives of these dimensionality reduction techniques are: 1) improvement in the accuracy of the multivariate approaches 2) reduction of the computational demand issue 3) offering a deeper and better understanding of the underlying processes (Mwangi et al. 2014). Most of the previous studies had focused on improving the first and second objectives (Hua et al. 2009; Franke et al. 2010; Wang et al. 2010; Chu et al. 2012). Until recently, there has been no detailed

investigation on the third objective. Thus, our project intended to improve the third objective without making any detrimental impact on the first and second objectives.

Recently, neuroimaging based studies focus greatly on multimodal setups. Through multimodal data analysis, integration of the complimentary information derived from different magnetic resonance imaging (MRI) measurements is possible, (Rykhlevskaia et al. 2008; Ritter et al. 2015; Miller et al. 2016; Liem et al. 2017). In this framework, implementation of the multivariate approaches on a multimodal setup improves the sensitivity and specificity of the outcomes, as multimodal data provide a more comprehensive representation of the brain than individual modalities (Erus et al. 2015; Davatzikos 2016; Liem et al. 2017). However, practical concatenation of multimodal data has to deal with several methodological limitations. Thus, the modern neuro-scientific community has spurred improvements within computation of multimodal neuroimaging, including pre-processing, feature extraction and fusion of the data. Given the dependency of the data fusion on the prior steps, our project is mainly dedicated for those prior steps, i.e. pre-processing and feature extraction (partly referring to the dimensionality reduction) at the individual modality level. Even though the information and the source of information differ between the modalities, some communal information is most likely shared among the modalities (Groves et al. 2012; Liu et al. 2015). Therefore, it is crucial to examine the latent dimensions of the data for each modality individually, prior to the investigation of patterns among the multimodal data obtained by linking different modalities together. In this framework, the, most frequently acquired modalities in big MRI data samples available for the research community are resting state functional MRI data and T1 based structural MRI data. Therefore, our studies focus on the influence of various data reduction techniques used in functional and structural MRI data.

Altogether, in this project, we mainly aimed to address the curse of dimensionality associated with high-dimensional voxel wise information by optimizing and evaluating various methods used in functional and structural MRI data separately, which can provide reliable and interpretable solutions. First part of the dissertation focused on functional MRI based data reduction methods, giving particular attention on reliability of the functional connectivity measures. Second part of the dissertation intended to examine the latent dimensionality of the structural MRI data, which predominantly improves the interpretability of the outcomes.

2.1 Reliability of the functional connectivity measures

When performing a multivariate approaches using functional MRI data, functional connectivity (FC) measure between different regions of the brain is considered to be

the dimensionality of variables. Nevertheless, voxel-wise whole brain functional connectivity profile provides a multitude of connectivity measures (i.e., 300000 x 300000), while only few of these connections are informative during the multivariate analysis of a given variable of interest (Wang 2011; Mwangi et al. 2014). In turn, a priori defined canonical networks yield considerably less connections, which in fact eliminate the influence of irrelevant features and retain the relevant ones for a given variable. Hence, a priori defined seed based FC measures are computationally more tractable and more informative. Therefore, our study 1 focused on region-to-region connectivity within *a priori* meta-analytically-defined networks (Wager et al. 2007; Schilbach et al. 2014).

Coordinates-based meta-analysis techniques are intended to reduce the features based on existing 'domain knowledge' (Dukart et al. 2013; Tench et al. 2013; Mwangi et al. 2014). The strength of this technique is to provide robust, functionally specific regions of interest (ROI) by integrating the outcomes from several different studies, which overcomes the limitation of false-positives ensuing due the low statistical power (i.e., smaller sample size) within a single study (Rottschy et al. 2012; Hardwick et al. 2013; Tench et al. 2013). Therefore, feature reduction based on meta-analysis techniques avoids the potentially poor reliability limitation encounter when selecting feature based on individual sample size study. However, the best strategy to extract the connectivity matrices from a priori networks still remains as a question. Impediment of the true measurement of neuronal activity from functional MRI signals caused by various nuisance signals leads to an unstable assessment of functional connectivity (Fox et al. 2005; Bright et al. 2017). Furthermore, the reliability of the functional connectivity might also be influenced based on different procedures of signal extraction from an ROI. Given the crucial importance of reliability for the development of clinical applications, study 1 aimed to identify the combination of signal extraction and confound removal approaches that yields the highest test-retest reliability when assessing resting-state functional connectivity in meta-analytically defined networks, using standard acquisitions as feasible in clinical practice. In study 1, we even investigated reliability from the two different but complementary perspectives that is, reliability at the subject level (RoSO) and reliability at the connection level (RoCO).

2.2 Feature reduction and feature selection

After addressing the best strategies to reliably extract the connectivity measures, the focus of our project is shifted towards evaluation of dimensionality reduction methods with a particular interest in reducing the high-dimensional feature space into biologically plausible reduced features. Feature reduction and feature selection are

two closely related terms in the context of multivariate approaches. As previously described, transformation of high dimensional voxel wise data into low rank approximations is considered as dimensionality reduction, also commonly known as feature reduction. In the feature reduction process, entire representation of the data is compressed into reduced number of features/components. The categorization of the data-driven feature reduction happens by inferring a hidden pattern of the data across the observations, which is also called as generative modeling or unsupervised learning of the data (Fukunaga 1990; Van Der Maaten et al. 2009). Nevertheless, inspite of an optimistically efficient feature reduction using data-driven methods, there still might be irrelevant or redundant features among the reduced set of features, influencing the performance of a multivariate approaches of a particular target variable. Hence, further selection of features that are relevant for that particular target variable is expected to diminish the influence of irrelevant or noisy features on the performance (Tipping and Faul 2003b; Hastie et al. 2015). This procedure of curbing the features contributed in the analysis is named as feature selection. With the recent developments in the field, sparse regression models allow us to perform a feature selection by learning the data in a supervised fashion or descriptive modeling. Therefore, second part of this dissertation (i.e., study 3 and study 4) was designed to evaluate the combination of sparse feature reduction procedure with a sparse supervised algorithm (to implement feature selection), aiming to provide solutions which can better interpret the patterns of brain regions contributed in the multivariate analysis. This is also called as Generative-descriptive method (Chu et al. 2012; Davatzikos 2016).

As introduced in the previous section, feature reduction technique depending on the existing domain knowledge (such as meta analytically derived networks) is an alternative procedure commonly employed for reducing the functional MRI data (Yarkoni et al. 2011; Dukart et al. 2013; Schilbach et al. 2014). Here, each node within a network is defined by consolidated evidence from multiple task-based fMRI studies. Thus, spatial correspondence of each node within the entire network reflects consistently co-activated brain region pertaining to the specific target behavioral condition across individuals (Laird 2009; Eickhoff et al. 2011). A priori defined meta analytical networks presumably indirectly induce sparsity to yield considerably less features among the whole brain connectivity by retaining the relevant features for the given target variable. Hence, study 2 and 3 attempted to implement the feature reduction based on domain knowledge during the investigation of the pattern implicated to the given variable using the functional MRI data. Here, study 2 aimed to examine the results obtained by employing an indirect sparse feature reduction procedure (i.e., meta analytically derived networks) to classify different neuropsychological disorders among each other's and also with the healthy controls

using a classical non-sparse regression model (i.e., support vector machine). In addition, study 3 focused on investigating the prediction performance of various meta analytically derived networks, when implementing a sparse regression model (which enforces feature selection procedure, i.e., relevance vector machine) to predict various personality traits.

Importantly, activations-based meta-analytic data representations are assumed to optimally summarize functional MRI data. In contrast, such representation might not be optimal for structural MRI data. Thus, in study 4 using structural MRI data, we investigate data representation based on structural modality. More concretely, study 4 strictly focused on evaluating data-driven techniques involved in reducing the dimensionality associated with the variables (i.e., voxel number) derived from voxel based morphometric (VBM) measure computed from anatomical MRI data. Importantly, while the results obtained from study 4 have been analyzed from an anatomical perspective of the brain, this technique may be applicable to other modalities. To note, our study 4 was performed on considerably large sample sizes to tackle the limitations associated with the statistical power of observations (between 700 to 1000 subjects). Given the interest for deeper understanding of the outcomes, study 4 attempted to implement an alternative dimensionality reduction method namely, non-negative matrix factorization (NNMF), which improves the quality of interpretability of the low-rank approximations in contrast to the frequently used principal component analysis method (Sotiras et al. 2015). NNMF based reduction of VBM data has been evaluated for prediction of age, using a sparse regression model (i.e., LASSO regularization model). To address the issue of computational demand, we examined the transferability of the NNMF between two independent datasets with different age distribution and acquisition protocols. Lastly but most importantly, as our project aimed to focus on the interpretability of the underlying processes, we investigated the association of the pattern of brain regions (i.e., reduced interpretable features) contributed in the regression analyses with the phenotype or behavioral score examined in each of the last three studies.

References

Bandettini PA, Wong EC, Hinks RS, et al (1992) Time course EPI of human brain function during task activation. Magn Reson Med 25:390–397. doi: 10.1002/mrm.1910250220

Boutsidis C, Gallopoulos E (2008) SVD based initialization: A head start for nonnegative matrix factorization. Pattern Recognit 41:1350–1362. doi: 10.1016/j.patcog.2007.09.010

Bright MG, Tench CR, Murphy K (2017) Potential pitfalls when denoising resting state fMRI data using nuisance regression. Neuroimage 154:159–168. doi:

10.1016/j.neuroimage.2016.12.027

Brown M, Semelka R, Nishino TK (2004) MRI: Basic Principles and Applications, 3rd edition.

Bunea F, She Y, Ombao H, et al (2011) Penalized least squares regression methods and applications to neuroimaging. Neuroimage 55:1519–1527. doi: 10.1016/j.neuroimage.2010.12.028

Button K (2014) Unreliable neuroscience ? Why power matters. Guard. 1–5.

Button KS, Ioannidis JP a, Mokrysz C, et al (2013) Power failure: why small sample size undermines the reliability of neuroscience. Nat Rev Neurosci 14:365–76. doi: 10.1038/nrn3475

Buxton RB (2002) Introduction to Functional Magnetic Resonance Imaging: Principles and Techniques. Energy 24:xi, 523 . doi: 10.1017/CBO9780511605505

Chang C, Lin C (2011) LIBSVM : A Library for Support Vector Machines. ACM Trans Intell Syst Technol 2:1–39. doi: 10.1145/1961189.1961199

Chu C, Hsu AL, Chou KH, et al (2012) Does feature selection improve classification accuracy? Impact of sample size and feature selection on classification using anatomical magnetic resonance images. Neuroimage 60:59–70. doi: 10.1016/j.neuroimage.2011.11.066

Cox DD, Savoy RL (2003) Functional magnetic resonance imaging (fMRI) "brain reading": Detecting and classifying distributed patterns of fMRI activity in human visual cortex. Neuroimage 19:261–270. doi: 10.1016/S1053-8119(03)00049-1

Craddock RC, Holtzheimer PE, Hu XP, Mayberg HS (2009) Disease state prediction from resting state functional connectivity. Magn Reson Med 62:1619–1628. doi: 10.1002/mrm.22159

Davatzikos C (2016) Computational neuroanatomy using brain deformations: From brain parcellation to multivariate pattern analysis and machine learning. Med. Image Anal. 33:149–154.

Davatzikos C, Bhatt P, Shaw LM, et al (2011) Prediction of MCI to AD conversion, via MRI, CSF biomarkers, and pattern classification. Neurobiol Aging. doi: 10.1016/j.neurobiolaging.2010.05.023

Dukart J, Mueller K, Barthel H, et al (2013) Meta-analysis based SVM classification enables accurate detection of Alzheimer's disease across different clinical centers using FDG-PET and MRI. Psychiatry Res - Neuroimaging 212:230–236. doi: 10.1016/j.pscychresns.2012.04.007

Eickhoff S, Laird A, Grefkes C, et al (2009) Coordinate-based ALE meta-analysis of neuroimaging data: a random-effects approach based on empirical estimates of spatial uncertainty. Hum Brain Mapp 30:2907–2926. doi: 10.1002/hbm.20718.Coordinate-based

Eickhoff SB, Bzdok D, Laird AR, et al (2012) Activation likelihood estimation meta-analysis revisited. Neuroimage 59:2349–2361. doi: 10.1016/j.neuroimage.2011.09.017

Eickhoff SB, Bzdok D, Laird AR, et al (2011) Co-activation patterns distinguish cortical modules, their connectivity and functional differentiation. Neuroimage 57:938–949. doi: 10.1016/j.neuroimage.2011.05.021

Erus G, Battapady H, Satterthwaite TD, et al (2015) Imaging patterns of brain development and their relationship to cognition. Cereb Cortex 25:1676–1684. doi: 10.1093/cercor/bht425

Fan Y, Shen D, Gur RC, et al (2007) COMPARE: Classification of morphological patterns using adaptive regional elements. IEEE Trans Med Imaging 26:93–105. doi: 10.1109/TMI.2006.886812

Fox MD, Snyder AZ, Vincent JL, et al (2005) The human brain is intrinsically organized into dynamic, anticorrelated functional networks. Proc Natl Acad Sci U S A 102:9673–8. doi: 10.1073/pnas.0504136102

Franke K, Ziegler G, Klöppel S, Gaser C (2010) Estimating the age of healthy subjects from T1-weighted MRI scans using kernel methods: Exploring the influence of various parameters. Neuroimage 50:883–892. doi: 10.1016/j.neuroimage.2010.01.005

Friedman J, Hastie T, Tibshirani R (2010) Regularization Paths for Generalized Linear Models via Coordinate Descent. J Stat Softw. doi: 10.18637/jss.v033.i01

Friston KJ (1997) Imaging cognitive anatomy. Trends Cogn Sci 1:21–7. doi: 10.1016/S1364-6613(97)01001-2

Friston KJ, Holmes AP, Worsley KJ, et al (1994) Statistical parametric maps in functional imaging: A general linear approach. Hum Brain Mapp 2:189–210. doi: 10.1002/hbm.460020402

Fukunaga K (1990) Introduction to Statistical Pattern Recognition.

Gonsalves BD, Cohen NJ (2010) Brain Imaging, Cognitive Processes, and Brain Networks. Perspect Psychol Sci 5:744–752. doi: 10.1177/1745691610388776

Good CD, Johnsrude IS, Ashburner J, et al (2001) A voxel-based morphometric study of ageing in 465 normal adult human brains. Neuroimage 14:21–36. doi: 10.1006/nimg.2001.0786

Groves AR, Smith SM, Fjell AM, et al (2012) Benefits of multi-modal fusion analysis on a large-scale dataset: Life-span patterns of inter-subject variability in cortical morphometry and white matter microstructure. Neuroimage 63:365–380. doi: 10.1016/j.neuroimage.2012.06.038

Guyon I, Elisseeff A (2003) An Introduction to Variable and Feature Selection. J Mach Learn Res 3:1157–1182. doi: 10.1016/j.aca.2011.07.027

Habeck CG (2010) Basics of Multivariate Analysis in Neuroimaging Data. J Vis Exp 1–6. doi: 10.3791/1988

Hardwick RM, Rottschy C, Miall RC, Eickhoff SB (2013) A quantitative meta-analysis and review of motor learning in the human brain. Neuroimage 67:283–297. doi: 10.1016/j.neuroimage.2012.11.020

Hastie T, Tibshirani R, Friedman J (2009) The Elements of Statistical Learning. Elements 1:337–387. doi: 10.1007/b94608

Hastie T, Tibshirani R, Wainwright M (2015) Statistical Learning with Sparsity: The Lasso and Generalizations. Crc 362. doi: 10.1201/b18401-1

Haxby J V, Gobbini MI, Furey ML, et al (2001) Distributed and Overlapping Representations of Face and Objects in Ventral Temporal Cortex. Science (80-) 293:2425–2430. doi: 10.1126/science.1063736

Horn, R. A. and Johnson CR (1973) Norms of Vectors and Matrices. Pure Appl Math 9:135–142. doi: 10.1016/S0079-8169(08)62705-1

Hua J, Tembe WD, Dougherty ER (2009) Performance of feature-selection methods in the classification of high-dimension data. Pattern Recognit 42:409–424. doi: 10.1016/j.patcog.2008.08.001

Jolliffe IT (2005) Principal component analysis. Appl Opt 44:6486. doi: 10.1007/SpringerReference_205537

Jolliffe IT (2002) Principal Component Analysis, Second Edition.

Kamitani Y, Tong F (2005) Decoding the visual and subjective contents of the human brain. Nat Neurosci 8:679–685. doi: 10.1038/nn1444

Laird AR (2009) ALE meta-analysis workflows via the BrainMap database: Progress towards a probabilistic functional brain atlas. Front Neuroinform. doi: 10.3389/neuro.11.023.2009

Laird AR, Fox PM, Price CJ, et al (2005) ALE meta-analysis: Controlling the false discovery rate and performing statistical contrasts. In: Human Brain Mapping. pp 155–164

Lee DD, Seung HS (1999) Learning the parts of objects by non-negative matrix factorization. Nature 401:788–91. doi: 10.1038/44565

Liem F, Varoquaux G, Kynast J, et al (2017) Predicting brain-age from multimodal imaging data captures cognitive impairment. Neuroimage 148:179–188. doi: 10.1016/j.neuroimage.2016.11.005

Liu S, Cai W, Liu S, et al (2015) Multimodal neuroimaging computing: a review of the applications in neuropsychiatric disorders. Brain Informatics 2:167–180. doi: 10.1007/s40708-015-0019-x

May A (2011) Experience-dependent structural plasticity in the adult human brain. Trends Cogn. Sci. 15:475–482.

Miller KL, Alfaro-Almagro F, Bangerter NK, et al (2016a) Multimodal population brain imaging in the UK Biobank prospective epidemiological study. Nat Neurosci 19:1523–1536. doi: 10.1038/nn.4393

Miller KL, Alfaro-Almagro F, Bangerter NK, et al (2016b) Multimodal population brain imaging in the UK Biobank prospective epidemiological study. Nat Neurosci 19:1523–1536. doi: 10.1038/nn.4393

Mourão-Miranda J, Bokde ALW, Born C, et al (2005) Classifying brain states and determining the discriminating activation patterns: Support Vector Machine on functional MRI data. Neuroimage 28:980–995. doi: 10.1016/j.neuroimage.2005.06.070

Mourão-Miranda J, Oliveira L, Ladouceur CD, et al (2012) Pattern recognition and functional neuroimaging help to discriminate healthy adolescents at risk for mood disorders from low risk adolescents. PLoS One. doi: 10.1371/journal.pone.0029482

Mukamel R, Gelbard H, Arieli A, et al (2005) Coupling between neuronal firing, field potentials, and fMR1 in human auditory cortex. Science (80-) 309:951–954. doi: 10.1126/science.1110913

Mwangi B, Tian TS, Soares JC (2014) A review of feature reduction techniques in Neuroimaging. Neuroinformatics 12:229–244.

O'Toole AJ, Jiang F, Abdi H, et al (2007) Theoretical, Statistical, and Practical Perspectives on Pattern-based Classification Approaches to the Analysis of Functional Neuroimaging Data. J Cogn Neurosci 19:1735–1752. doi:

10.1162/jocn.2007.19.11.1735

Ogawa S, Tank DW, Menon R, et al (1992) Intrinsic signal changes accompanying sensory stimulation: Functional brain mapping with magnetic resonance imaging. Proc Natl Acad Sci U S A 89:5951–5955. doi: 10.1073/pnas.89.13.5951

Pereira F, Mitchell T, Botvinick M (2009) Machine learning classi!ers and fMRI: A tutorial overview. Neuroimage 45:S199–S209. doi: 10.1016/j.neuroimage.2008.11.007.Machine

Ritter K, Schumacher J, Weygandt M, et al (2015) Multimodal prediction of conversion to Alzheimer's disease based on incomplete biomarkers∗This work was supported by the Bernstein Computational Program of the German Federal Ministry of Education and Research (01GQ1001C, 01GQ0851, GRK 1589/1), the Europe. Alzheimer's Dement Diagnosis, Assess Dis Monit 1:206–215. doi: 10.1016/j.dadm.2015.01.006

Rottschy C, Langner R, Dogan I, et al (2012) Modelling neural correlates of working memory: A coordinate-based meta-analysis. Neuroimage 60:830–846. doi: 10.1016/j.neuroimage.2011.11.050

Rykhlevskaia E, Gratton G, Fabiani M (2008) Combining structural and functional neuroimaging data for studying brain connectivity: A review. Psychophysiology 45:173–187.

Schilbach L, Müller VI, Hoffstaedter F, et al (2014) Meta-analytically informed network analysis of resting state fMRI reveals hyperconnectivity in an introspective socio-affective network in depression. PLoS One. doi: 10.1371/journal.pone.0094973

Schölvinck ML, Maier A, Ye FQ, et al (2010) Neural basis of global resting-state fMRI activity. Proc Natl Acad Sci U S A 107:10238–43. doi: 10.1073/pnas.0913110107

Schrouff J, Rosa MJ, Rondina JM, et al (2013) PRoNTo: Pattern recognition for neuroimaging toolbox. Neuroinformatics 11:319–337. doi: 10.1007/s12021-013-9178-1

Scott A. Huettel; Allen W. Song; Gregory McCarthy, Huettel, Scott; Song A, Beckmann CF, et al (2004) Functional Magnetic Resonance Imaging.

Sotiras A, Resnick SM, Davatzikos C (2015) Finding imaging patterns of structural covariance via Non-Negative Matrix Factorization. Neuroimage 108:1–16. doi: 10.1016/j.neuroimage.2014.11.045

Symms M, Jäger HR, Schmierer K, Yousry T a (2004) A review of structural magnetic resonance neuroimaging. J Neurol Neurosurg Psychiatry 75:1235–1244. doi: 10.1136/jnnp.2003.032714

Tench CR, Tanasescu R, Auer DP, Constantinescu CS (2013) Coordinate Based Meta-Analysis of Functional Neuroimaging Data; False Discovery Control and Diagnostics. PLoS One 8:1–12. doi: 10.1371/journal.pone.0070143

Tibshirani R (1996) Regression Selection and Shrinkage via the Lasso. J. R. Stat. Soc. B 58:267–288.

Tipping M (2001) Sparse Bayesian Learning and the Relevance Vector Mach. J Mach Learn Res 1:211–244. doi: 10.1162/15324430152748236

Tipping ME, Faul AC (2003a) Fast marginal likelihood maximisation for sparse Bayesian models. In: Proceedings of the ninth …. pp 1–13

Tipping ME, Faul AC (2003b) Fast Marginal Likelihood Maximisation for Sparse Bayesian Models. Ninth Int Work Aritficial Intell Stat 1–13. doi: 10.1.1.165.4281

Tisserand DJ, Pruessner JC, Sanz Arigita EJ, et al (2002) Regional frontal cortical volumes decrease differentially in aging: An MRI study to compare volumetric approaches and voxel-based morphometry. Neuroimage 17:657–669. doi: 10.1016/S1053-8119(02)91173-0

Turkeltaub PE, Eden GF, Jones KM, Zeffiro TA (2002) Meta-analysis of the functional neuroanatomy of single-word reading: Method and validation. Neuroimage 16:765–780. doi: 10.1006/nimg.2002.1131

Van Der Maaten LJP, Postma EO, Van Den Herik HJ (2009) Dimensionality Reduction: A Comparative Review. J Mach Learn Res 10:1–41. doi: 10.1080/13506280444000102

Vapnik VN (1998) Statistical Learning Theory. Adapt Learn Syst Signal Process Commun Control 2:1–740. doi: 10.2307/1271368

Wager TD, Lindquist M, Kaplan L (2007) Meta-analysis of functional neuroimaging data: Current and future directions. Soc Cogn Affect Neurosci 2:150–158. doi: 10.1093/scan/nsm015

Wang Y (2011) High dimensional pattern regression. Image (Rochester, NY) 50:1519–1535. doi: 10.1016/j.neuroimage.2009.12.092.High-Dimensional

Wang Y, Fan Y, Bhatt P, Davatzikos C (2010) High-dimensional pattern regression using machine learning: From medical images to continuous clinical variables. Neuroimage 50:1519–1535. doi: 10.1016/j.neuroimage.2009.12.092

Weishaupt D, Kochli VD, Marincek B, Kim EE (2007) How Does MRI Work? An Introduction to the Physics and Function of Magnetic Resonance Imaging. J Nucl Med 48:1910–1910. doi: 10.2967/jnumed.107.045104

Yang Z, Oja E (2010a) Linear and nonlinear projective nonnegative matrix factorization. IEEE Trans Neural Networks 21:734–749. doi: 10.1109/TNN.2010.2041361

Yang Z, Oja E (2010b) Linear and Nonlinear Projective Nonnegative Matrix Factorization. Trans Neur Netw 21:734–749. doi: 10.1109/TNN.2010.2041361

Yang Z, Yuan Z, Laaksonen J (2007) Projective Non-negative matrix factorization with application to facial image processing. Pattern Recognit 21:1353–1362.

Yarkoni T, Poldrack RA, Nichols TE, et al (2011) Large-scale automated synthesis of human functional neuroimaging data. Nat Methods 8:665–670. doi: 10.1038/nmeth.1635

Yoon JH, Tamir D, Minzenberg MJ, et al (2008) Multivariate Pattern Analysis of Functional Magnetic Resonance Imaging Data Reveals Deficits in Distributed Representations in Schizophrenia. Biol Psychiatry 64:1035–1041. doi: 10.1016/j.biopsych.2008.07.025

Zou H, Hastie T (2005) Regularization and variable selection via the elastic-net. J R Stat Soc 67:301–320. doi: 10.1111/j.1467-9868.2005.00503.x

STUDY 1

Resting-state test-retest reliability of *a priori* defined canonical networks over different preprocessing steps

Deepthi P. Varikuti[1,2,], Felix Hoffstaedter [1,2,], Sarah Genon[1,2], Holger Schwender[3], Andrew T. Reid[2], Simon B. Eickhoff [1,2]

[1]*Institute of Clinical Neuroscience and Medical Psychology, Medical Faculty, Heinrich Heine University Düsseldorf, Germany*
[2]*Institute of Neuroscience and Medicine (INM-1), Research Centre Juelich, Germany*
[3]*Mathematical Institute, Heinrich Heine University Düsseldorf, 40225 Düsseldorf, Germany.*

Own contributions	Conception and design of experiment
	Reviewing and adapting analysis code
	Statistical data analysis
	Interpretation of results
	Preparing figures
	Writing the paper
	Total contribution 80%

CrossMark

ORIGINAL ARTICLE

Resting-state test–retest reliability of a priori defined canonical networks over different preprocessing steps

Deepthi P. Varikuti[1,2] · Felix Hoffstaedter[1,2] · Sarah Genon[1,2] · Holger Schwender[3] ·
Andrew T. Reid[2] · Simon B. Eickhoff[1,2]

Abstract Resting-state functional connectivity analysis has become a widely used method for the investigation of human brain connectivity and pathology. The measurement of neuronal activity by functional MRI, however, is impeded by various nuisance signals that reduce the stability of functional connectivity. Several methods exist to address this predicament, but little consensus has yet been reached on the most appropriate approach. Given the crucial importance of reliability for the development of clinical applications, we here investigated the effect of various confound removal approaches on the test–retest reliability of functional-connectivity estimates in two previously defined functional brain networks. Our results showed that gray matter masking improved the reliability of connectivity estimates, whereas denoising based on principal components analysis reduced it. We additionally observed that refraining from using any correction for global signals provided the best test–retest reliability, but failed to reproduce anti-correlations between what have been previously described as antagonistic networks. This suggests that improved reliability can come at the expense of

potentially poorer biological validity. Consistent with this, we observed that reliability was proportional to the retained variance, which presumably included structured noise, such as reliable nuisance signals (for instance, noise induced by cardiac processes). We conclude that compromises are necessary between maximizing test–retest reliability and removing variance that may be attributable to non-neuronal sources.

Introduction

Functional magnetic resonance imaging (fMRI) relies on the measurement of changes in blood oxygenation (i.e., BOLD) and plays a vital role in understanding normal and abnormal brain functioning. For instance, functional connectivity of distant brain regions can be investigated through the statistical analysis of coherent low-frequency BOLD fluctuations. Synchronized signal fluctuations can be observed even when the subject is at rest, without performing any task, and the analysis of resting-state data has become a popular means of studying ongoing brain activations and functional connectivity between brain regions (Biswal et al. 1995; Fox and Raichle 2007). There are both indirect (from comparison with task co-activation patterns, (Kwong et al. 1992; Hinke et al. 1993; Buckner et al. 1996; Huettel et al. 2004; Barch et al. 2013)) and direct (from invasive recordings (He et al. 1999; Lai et al. 2011; Lu et al. 2014) supports toward this notion. Several confounding effects, including system noise, thermal noise, and noise induced by non-neuronal physiological processes, may influence the measured signal and hence

✉ Deepthi P. Varikuti
d.varikuti@fz-juelich.de

[1] Institute of Clinical Neuroscience and Medical Psychology, Medical Faculty, Heinrich Heine University Düsseldorf, Düsseldorf, Germany

[2] Institute of Neuroscience and Medicine (INM-1), Research Center Juelich, 52425 Juelich, Germany

[3] Mathematical Institute, Heinrich Heine University Düsseldorf, 40225 Düsseldorf, Germany

Springer

apparent brain activity. Therefore, interpretation of brain activity depends on the ability to mitigate their influences (Fox et al. 2009).

Participant-induced artifacts, such as motion and physiologically induced artifacts (i.e., due to respiration and cardiac processes), comprise the largest component of noise affecting the BOLD signal (Windischberger et al. 2002). Motion artifacts have been shown to produce spurious correlations in a systematic way (Van Dijk et al. 2012; Power et al. 2012; Satterthwaite et al. 2013), implying that the removal of motion related artifacts is a prerequisite for further analysis. Various approaches have been proposed for dealing with noise effects post hoc, i.e., after the data have been acquired (Behzadi et al. 2007; Fox et al. 2009; Murphy et al. 2009; Chai et al. 2012; Griffanti et al. 2014; Patriat et al. 2015; Power et al. 2015; Soltysik et al. 2015; Wong et al. 2016). In addition to motion-related artifacts, one particular aspect that has received a lot of attention is the use of nuisance regressors reflecting global signals, derived either from the whole brain or from specific tissue types, such as white-matter or cerebrospinal fluid. However, removal of various global nuisance regressors alters the variance of the residual signal and has been shown to modify the correlational structure (Fox et al. 2009). In line with it, Friston (2011) showed that changing the signal-to-noise ratio can change the correlation coefficient, which indicates that the level of observable noise influences the correlation coefficient.

Both the definition of the ROI from which BOLD signals are extracted and the means by which voxel-wise signals are summarized across a given ROI are critical considerations in a functional connectivity analysis. An ROI can be derived through various approaches, including (most simply) a single voxel or sphere of a fixed radius around a voxel, histological parcellation in standard space (Eickhoff et al. 2005), clustering approaches based on functional or structural connectivity estimates (Eickhoff et al. 2015), thresholded statistical maps, or meta-analytic approaches such as ALE (Eickhoff et al. 2009, 2012). In this study, we focused on region-to-region connectivity within a priori meta-analytically defined networks (Schilbach et al. 2014; Schilbach 2016). This approach has several advantages. In particular, meta-analyses provide robust, functionally specific ROIs based on observations across many studies. Analyzing functional connectivity on this network combines its functional specificity with the advantages of task-free imaging, i.e., an acquisition that poses little demands on the subjects and is not confounded by a specific task paradigm. Similarly, the extraction of a summary signal across an ROI can be performed in various ways that may impact the reliability of connectivity estimates. In particular, the exclusion of voxels based on their gray matter probabilities may help improve signal to noise

by removing signal not originating in the gray matter tissue of interest. In this study, we compared three signal extraction approaches using different gray matter masking techniques.

Many clinical studies currently rely on functional connectivity measures in understanding normal and abnormal brain functioning. The appeal for resting-state functional connectivity analyses in clinical applications lies in the fact that such data are easy to acquire without any specific setup, do not require active participation by the subjects, and in contrast to task-based data, are less influenced by compliance and performance. Nevertheless, several concerns have been raised regarding the reproducibility and statistical power of classical neuroimaging studies (Button et al. 2013a, b). Clinical application, however, can only be useful if the analyses yield reliable measures. Various studies have also been performed to test the reliability of functional or effective connectivity measures using different modalities (such as fMRI or diffusion MRI) and reported moderate to high test–retest reliability of connectivity measures across moderate to long-term scans. (Chen et al. 2015; Frässle et al. 2016; Song et al. 2016; Zhong et al. 2015). Shehzad et al. (2009) investigated the test–retest reliability of global connectivity patterns using resting-state fMRI and observed that significant connectivity scores are more reliable than non-significant connectivity scores. Wang et al. (2011) evaluated short-term (less than 1 h apart) and long-term (more than 5 months apart) test–retest reliability for topological metrics of functional networks and observed that long-term scans had better reliability than short-term scans. Later, Raemaekers et al. (2012) analyzed the reliability of BOLD activation and reported that patterns of BOLD activation were relatively stable across sessions, while the amplitude of the activations is more variable. Gorgolewski et al. (2013) studied the test–retest reliability of confound removal at the subject level (by focusing on the single subject reliability) and showed that subject motion can detrimentally impact reliability. Yan et al. (2013b) investigated the influence of post-acquisition standardization techniques on traditional fMRI measures, test–retest reliability, and phenotypic relationships, as well as nuisance variables (mainly mean global signal) and reported that global signal regression is identical to gray matter regression and both should be avoided. Subsequently, Birn et al. (2014) evaluated the influence of various physiological noise correction methods on test–retest reliability and found that it was reduced by physiological noise correction, as it reduced the variability between subjects as well as within the subject. Shirer et al. (2015) investigated various means of confound removal across multiple outcome measures and demonstrated that noisiness, reliability, and heterogeneity of the data varies based on the preprocessing parameter chosen. In turn, the

influence of various gray matter masking approaches on the reliability has not been addressed in any of the previous test–retest studies. Therefore, using meta-analytically derived networks, we assessed the influence of different signal extraction and noise regression approaches on the reliability of the resting-state functional connectivity measures.

In this study, we evaluated test–retest reliability of resting-state functional connectivity in a cohort of 42 subjects scanned twice with a between-scan interval of 175 ± 75 days. We assessed two networks: the extended socio-affective default mode network (eSAD) (Amft et al. 2015); and the working-memory network (WMN) (Rottschy et al. 2012). Both networks were derived from previous meta-analytic studies, which used anatomical likelihood estimation (ALE; (Eickhoff et al. 2009, 2012)) to identify regions that are robustly activated across studies, for specific task paradigms. Both networks have been hypothesized to anti-correlate with each other (Fox et al. 2005; Reid et al. 2016). Thus, the reliability of connectivity estimates within, as well as between, the specified meta-analytically derived networks was evaluated.

A literature survey was conducted, to investigate the popularity of various methods for confound removal in recent fMRI studies. Using PubMed database, all the articles with the terms "fMRI," "resting-state," and "seed-based," published from the beginning of 2014 until the time of this study (June 2016) were identified, reflecting the recent work most in line with the focus of our work on seed-based analyses. A total number of 239 studies were identified. Among them, 33 studies had to be excluded, because the articles were either not relevant to the study (such as studies on animals) or not accessible. Therefore, a total number of 206 studies were investigated. We then computed the percentage of studies using the different confounds removal methods, which is shown in Fig. 1. The frequency of studies when using a certain confound has been demonstrated separately (in the categories of 'Only') and in combination with the other confounds in the Fig. 1.

Based on this literature examination, we assessed the effects of the most commonly used confound removal approaches in resting state fMRI studies; namely, global and tissue-class specific (either only WM and CSF or in addition also GM) mean signal regression, as well as
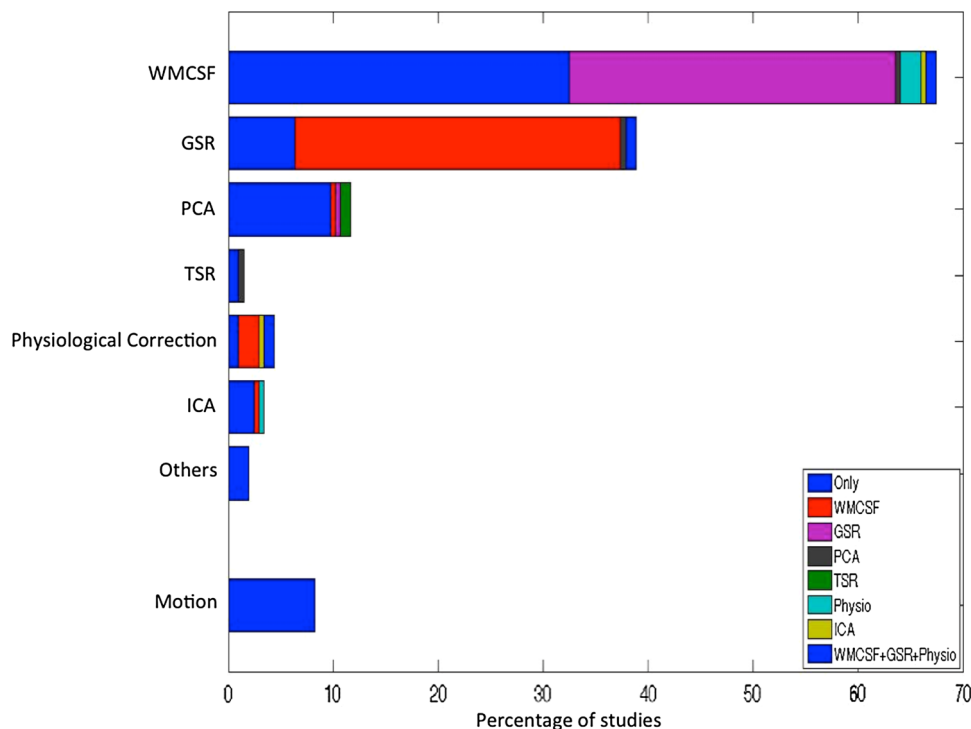
Fig. 1 Percentage of studies using a certain confound removal method [i.e., white matter and cerebral spinal fluid signal regression (WMCSF), global signal regression (GSR), principle component analysis-based corrections (PCA), tissue signal regression (TSR), physiological recordings-based corrections (physiological correction), independent component analysis-based corrections (ICA), and other correction methods such as ANATICOR or gray matter atrophy regression (others)]. The *colors* represent the interactions of each method with other methods. The first fraction of section which is consistent over the approaches, represented with the word "Only" (in *blue*) shows the percentage of studies performing only a certain confound removal without any interactions. Additional colors assigned to the other confound removal appears only when there is an interaction. Of note, the interactions of motion regression with other methods are not explicitly shown in this figure. However, almost all the studies involved in this literature survey have removed the motion effects along with the other confound removal approaches demonstrated in the figure

principal components analysis (PCA) denoising. We also examined three approaches for extracting the regional time-series based on different methods for gray matter masking. Above mentioned approaches were assessed separately and in combination with each other. To observe the consequences of the interactions, the approaches were evaluated in combinations (cf. "Assessed (combinations of) signal processing steps"). We note that physiological noise regression (i.e., elimination of artifacts induced by respiration and cardiac processes) requires recordings of parameters, such as heartbeat and breathing. Such physiological recordings, however, are rarely acquired in standard (clinical) resting-state acquisitions and were hence not considered in the current investigation. Independent component analysis (ICA)-based denoising is another emerging approach to confound removal (Griffanti et al. 2014; Salimi-Khorshidi et al. 2014; Pruim et al. 2015a, b). However, ICA-based denoising approaches (excluding the ICA-AROMA, as the pre-defined spatial features included within in the package itself) require effective individual segmentation from high-resolution T1 images, which were not available for the current data. Acknowledging the future potential of ICA-based denoising, we thus focused our work on the evaluation of the presently most widely used approaches.

Another common application of ICA is the examination of the functional connectivity networks. Recently, such ICA method followed with the dual regression is used to assess the functional connectivity for group comparisons, instead of seed-based functional connectivity. Zuo et al. (2010) reported moderate-to-high test–retest reliability. Furthermore, Smith et al. (2014) claimed that ICA followed with the dual regression performs better than the seed-based connectivity measures. Even though, such methods may lead to higher reliability. Zuo et al. (2010) reported moderate-to-high test–retest reliability, while computing the functional connectivity networks using ICA combined with the dual regression. Furthermore, Smith et al. (2014) investigated that ICA followed with the dual regression performs better than the seed-based connectivity measures. Even though such methods may lead to higher reliability (Zuo et al. 2010), seed-based functional connectivity is still very widely used for the examination of a priori hypotheses in both basic and clinical studies (Smith et al. 2014). Thus, we here focused on the test–retest reliability of the seed-based functional connectivity measures.

Importantly, reliability can be examined from two perspectives: at the subject level and at the connection level. One the one hand, meaningful group comparisons largely depend on reliability at the subject level, i.e., over scans the order of subjects should remain as similar as possible for any given connection. One the other hand, network modeling capitalizing on within-subject connectivity requires reliability at the connection level (cf. "Indices of reliability"), i.e., for any given subjects, the order of connectivity strengths should remain as similar as possible over scans. Therefore, in this study, we investigated reliability from the two different but complementary perspectives, that is, reliability at the subject level (RoSO) and reliability at the connection level (RoCO). To sum up, this study aimed to identify the combination of signal extraction and confound removal approaches that yield the highest test–retest reliability when assessing resting-state functional connectivity in meta-analytically defined networks, using standard acquisitions as feasible in clinical practice. In other words, this study aims to provide a ranking of methods in terms of their potential to yield stable connectivity patterns over time.

Materials and methods

Networks of interest

The influence of different processing steps on the test–retest reliability of resting-state functional connectivity analyses was assessed in two canonical networks related to cognitive and socio-affective processing. In particular, the two networks were defined by large-scale synthesis of neuroimaging findings using coordinate-based meta-analyses (Fox et al. 2014). As a prototypical "task-positive" cognitive network (regions exhibiting increase in activity during task performance), we assessed the core working memory network (WMN) described by Rottschy et al. (2012), consisting of nine bilateral fronto-parietal regions (Fig. 2a; Table 1). As a "task-negative" network (regions exhibiting decrease in activity during task performance), we included the extended socio-affective default mode (eSAD) network identified by Amft et al. (2015), which extended a previous meta-analytical definition of the default mode (Schilbach et al. 2012) and comprised 12 regions mainly corresponding to cortical midline structures (Fig. 2b; Table 1). Importantly, both of these networks have shown a strong positive coupling among their respective nodes but were anti-correlated with each other. They may thus be considered as robustly a priori defined network models for the often-proposed large-scale anti-correlated systems in the human brain (Fox et al. 2005).

Sample characteristics, preprocessing and RS-FC computation

Images acquisition

Resting-state fMRI data of 42 healthy subjects including 19 females with an average age of 42 ± 20 (mean $\pm$ std)
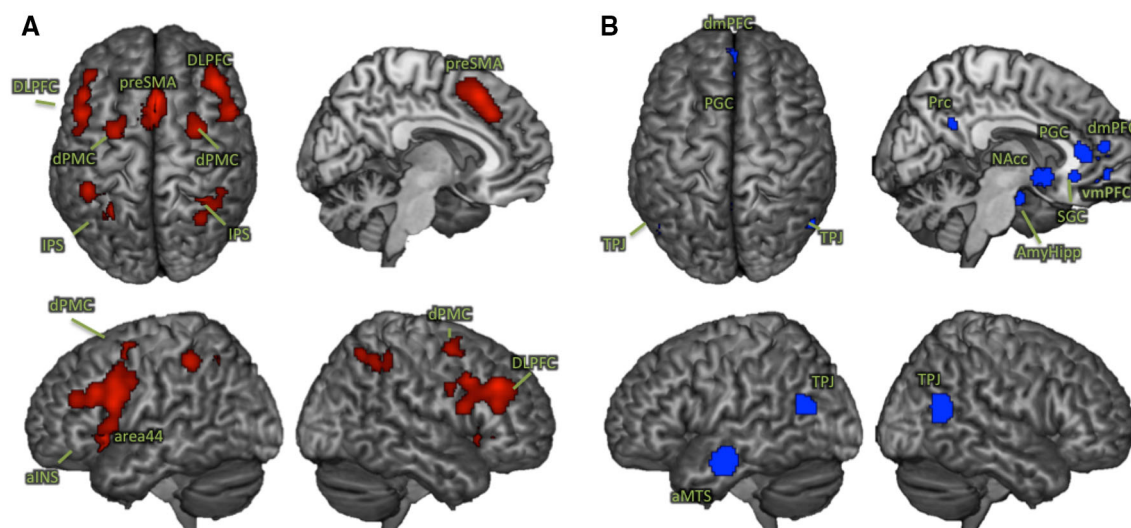
Fig. 2 Nodes of meta-analytically derived networks used for the reliability assessment. **a** The core working memory network (Rottschy et al. 2012). **b** The extended socio-affective default mode network (Amft et al. 2015)

Table 1 Coordinate details and cluster size (k) of the regions, within the WMN and eSAD involved in this study

	Macro-anatomical labels	Abbreviation	Side	k (voxels of size 3.1 mm isotropic)	MNI coordinates in standard RAS orientation		
					X	Y	Z
Working memory network (WMN) nodes							
1	Anterior insula	aINS	L	276	32	22	−4
2			R	182	34	28	−2
3	Dorsolateral prefrontal cortex	DLPFC	L	1331	50	12	22
4			R	1032	44	34	32
5	Pre-supplementary motor area	preSMA		1035	2	20	50
6	Intraparietal sulcus	IPS	L	543	30	56	48
7			R	310	36	48	44
8	Dorsal premotor cortex	dPMC	L	190	28	0	58
9			R	243	30	2	56
Extended socio-affective default mode (eSAD) network nodes							
10	Pregenual anterior cingulate cortex	ACC		180	0	36	10
11	Anterior middle temporal sulcus	aMTS	L	468	54	10	−20
12	Amygdala/hippocampus	Amy/hippo	L	86	24.0	10	−20
13			R	141	24	−8.0	−22.0
14	Basal ganglia	BG	L	146	−6	10	−8
15			R	188	6	10	−8
16	Dorsomedial prefrontal cortex	dmPFC		204	−2	52	14
17	Precuneus	PrC		145	−2	52	26
18	Subgenual anterior cingulate cortex	sACC		244	−2	32	−8
19	Temporo-parietal junction	TPJ	L	251	46	66	18
20			R	373	50	60	18
21	Ventromedial prefrontal cortex	vmPFC		114	−2	50	−10

years were obtained in two sessions with an average time interval of 175 ± 75 (mean ± std) days. In each session, 250 resting state EPI images were obtained on a Siemens 3T Scanner (Scanning parameters: TR 2200 ms, TE 30 ms, flip angle 90°, 36 slices, a voxel size 3.1 mm isotropic) corresponding to a scanning time of 9.2 min, which stays well in line with the reliable intersession scanning time of 8–12 min suggested by Birn et al. (2013). High-resolution

T1-weighted structural images were not acquired for the data set used in this study. The original study protocol of the data used here has been approved by the local ethics committees of the university hospital Aachen, and informed consent was obtained by all the participants prior to the examination. The current data were analyzed anonymously.

Images preprocessing

Prior to further processing (using SPM8, http://www.fil.ion.ucl.ac.uk/spm), the first four images were discarded allowing for magnetic field saturation. The EPI images were corrected for head movement by affine registration using a two-pass procedure. In a two-pass procedure, all the EPI images were aligned to the first EPI image. Then, a mean over the aligned EPI images was computed. Finally, all the EPI images were again aligned to the first pass mean EPI image. The mean EPI image for each subject was non-linearly normalized to the MNI152 non-linear template space template using the "unified segmentation" approach (Ashburner and Friston 2005). The ensuing deformation field was applied to the individual EPI volumes and smoothed with a 5-mm FWHM Gaussian kernel. Preprocessed images were visually checked for any processing artifacts.

Each node of the assessed functionally defined networks (cf. "Networks of interest") available in the same space was represented by its peak's coordinate. The time series for all voxels within a priori meta-analytically derived clusters were then extracted. Following gray matter masking if applicable (cf. "Assessed (combinations of) signal processing steps"), we then employed a multiple regression approach to control for different confounds in the EPI time series. While the choices for dealing with global signals were outlined below, we always included the six motion parameters derived from the image realignment as well as their derivative as the first-order (linear) and second-order (quadratic) terms as evaluated by (Satterthwaite et al. 2013). That is, in addition to the approach-specific confounds, these 24 movement regressors were used in all analyses. Following the removal of any variance in the individual voxels' time series that could be explained by the respective confounds, the data were band pass filtered preserving BOLD frequencies between 0.01 and 0.08 Hz (Biswal et al. 1995; Fox and Raichle 2007). We computed the frame-to-frame differences from the six motion parameters derived from the image realignment to assess frame-wise displacements (FD). An FD threshold of 0.5 mm was used to discard potentially motion-contaminated images, before bandpass filtering (Power et al. 2012; Yan et al. 2013a). Finally, the characteristic time series of each seed was computed as the first eigenvariate of the preprocessed time series for the individual voxels within that seed. The functional connectivity between every pair of nodes was then computed as the correlation coefficient between these time series, which were transformed to Fischer's Z scores to render them normally distributed (Fig. 3). Here, in this study, tissue class segmentation is performed on a mean EPI volume due to the lack of high-resolution T1 structural scans. Nevertheless, the registration of EPI images to T1 structural scans may fail to detect the non-linear distortions of the EPI images, especially in the absence of the field maps or such relevant images. However, partial volume effects may exist in the mean EPI volume based segmentation. To avoid such partial volume effects, gray matter masking along with a median-split approach, which extracts the signal only from 50 % of the voxels with high gray matter probability, has been implemented and evaluated in this study. In addition, median-split approach has an advantage of accounting similar number of voxels while extracting the signal, particularly when using meta-analytically derived clusters.

Assessed (combinations of) signal processing steps

As the key aim of this study was to assess the impact of different commonly used processing steps on the reliability of RS-FC measurements, we focused on three different domains as follows.

(I) Extraction of time series: Evidently, meaningful signal should mainly be found in gray matter (GM). Hence, the voxels within 5 mm of the seed's coordinate might be anatomically constrained based on tissue class segmentation as provided by SPM (Ashburner and Friston 2005). Here, we evaluated three options:

No gray matter mask (*NoGM*) All voxels within 5 mm of the seed coordinate were included, processed by confound removal and temporal filtering, and summarized by their first eigenvariate. No gray matter masking is the most commonly used approach in RS-FC analysis. Conceptually, NoGM considered the influence of cortical anatomy as minor relative to the spread of BOLD data and spatial smoothing.

Individual gray matter mask (*IndGM*) The GM probability as estimated by the unified segmentation for that particular subject was extracted for each voxel within 5 mm of the seed coordinate and a median-split approach was then performed retaining those 50 % of voxels with highest GM probabilities. This approach was based on the argument that the individual anatomy should be most important for tissue classification.

Group gray matter mask (*GrpGM*) The tissue class segmentations of all individual subjects were first averaged and a median-split approach of the voxels was then performed based on these average GM probabilities. In this

Fig. 3 Pipeline of the entire preprocessing steps until the RS-FC computation: the assessed combinations (inside the *red dotted box*) indicate the signal processing methods for which the reliability is evaluated in three different domains ((I) extraction of time series, (II) PCA denoising, (III) global signal removal)
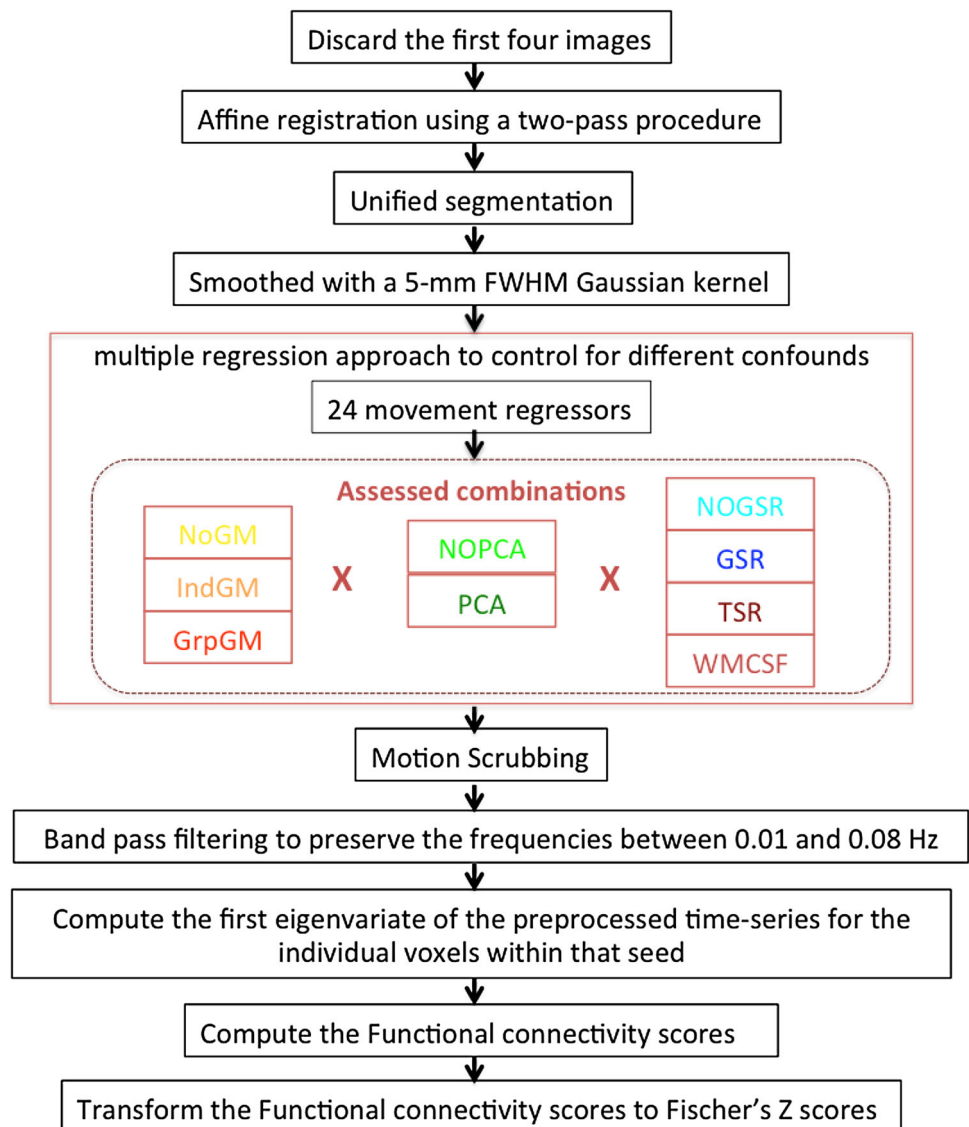
method, the focus on GM was retained but rather than basing the masking on the (prominently noisy) individual segmentation, group data (considered as more robust) were used. For reader's information, the overlap between the *IndGM* and *GrpGM* was computed and is shown in Fig. 4.

(II) PCA denoising: It has been suggested (Behzadi et al. 2007; Soltysik et al. 2015) that computing a principal component analysis (PCA) decomposition across the WM and CSF regions of the brain and removing variance associated with the most dominant five components might remove many sources of artificial and confounding signals and hence increase the specificity of RS-FC results. We thus performed all analyses both with (PCA) and without (NoPCA) PCA denoising.

(III) Global signal removal: As removing the global signal had received a lot of attention in recent discussions (Murphy et al. 2009; Chai et al. 2012; Saad et al. 2012; Fox

et al. 2013), we assessed seven different methods for this particular aspect. In that context, tissue class-specific global signals were computed based on the SPM8 segmentation of the (mean) EPI into GM, WM, and CSF regions, then averaged the signal time series of the voxels specific to each tissue class.

Global signal regression (*GSR*) Removes all variance explained by the first-order effects of the global (average across all voxels at each time point) signal.

Tissue signal regression (*TSR*) Removes variance explained by the first-order effects of the mean GM, WM, and CSF signals.

WM and CSF signal regression (*WMCSF*) The mean signals of the WM and CSF were removed, i.e., only the first-order effects.

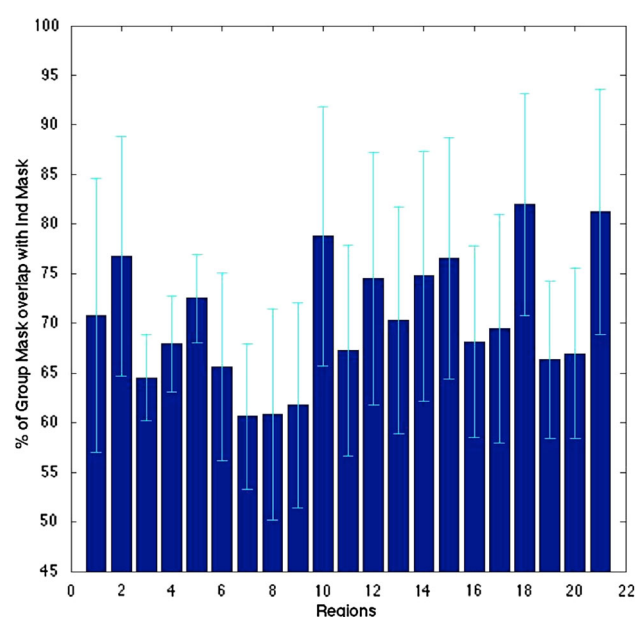*No global signal regression* (*NoGSR*) No removal of any global signal.

Fig. 4 Percentage of voxels that overlap between the individual and group masks, relative to the *GrpGM* for each of the 21 seed regions

Importantly, the different choices for each of the three main factors may be implemented independently of the other factors, allowing for a full permutation of the different analyses options and hence 42 (3 × 2 × 7) different combinations for signal extraction and confound removal. We, therefore, performed reliability analysis for all of these 42 combinations, i.e., analytical approaches.

Indices of reliability

To quantify the test–retest reliability of the 42 different approaches, we used two complementary measures that were each applied from two different perspectives. Test–retest reliabilities are quite often assessed using intra-class correlations (ICC), which takes into account inter-subjects variability in relation with the intra-subject variability. The intention of our study, however, was to examine one effect at a time, i.e., to evaluate inter-subject variability (i.e., RoSO) separately from intra-subject variability (i.e., RoCO). Therefore, reliability was tested using two measures. The first employed measure was Kendall's rank correlation (to quantify the consistency in relative order; Zang et al. 2004; Shehzad et al. 2009; Guo et al. 2011; Thomason et al. 2011; Li et al. 2012; Patriat et al. 2013) between the functional connectivity scores obtained at the first and second sessions, which quantifies the degree to which the order of observations is similar across both sessions. Modifications in the signal extraction and confound removal methods alter the residual signal fluctuations, which lead to variation in the connectivity measures. Thus, the stability of the relative orders when comparing

different connections/subjects was measured using Kendall's correlations. Complementing this index, we computed the absolute difference between functional connectivity scores to probe the numerical test–retest reliability. This index should be less sensitive to single outliers, in comparison with other alternatives like sum of squared measures. Thus, numerical differences when comparing different connections/subjects were measured using mean absolute differences.

These indices were computed from two different perspectives, reflecting the reliability at the subject level and at the connection level, respectively. In that context, reliability at the connection level (RoCO) addresses the question "are, for a given subject, the connections in the same order across sessions?" which was a prerequisite for any within-subject network modeling. We thus computed for each subject the correlation (across connections) between the first and second sessions (Fig. 5a) as well as the absolute difference between the two sessions by averaging them over connections (Fig. 5c). This perspective thus yields for every approach as many data points as there were subjects' within/between the two networks. Reliability at the subject level (RoSO) addresses the question "are, for a given connection, the subjects in the same order across sessions?" which was a prerequisite for group comparisons. Here, we computed for each connection the correlation (across subjects) between the first and second sessions (Fig. 5b) as well as the absolute differences between the two sessions by averaging them over subjects (Fig. 5d). This perspective thus yields, for every approach, the same number of data points, as there are connections in the respective network.

Finally, we computed two further important parameters in addition to these indices of reliability. First, the amount of variance within the extracted time series at the two time points was computed for each combination of methods to quantify the influence of confound removal on the variance of the residual resting-state signal (Fig. 5e). Second, for every approach, we computed percentage of positive connectivity scores among within-network (i.e., within eSAD and WMN regions) and between-network connections (i.e., between eSAD and WMN regions).

Aggregation and evaluation

The 42 different approaches defined by the combination of different masking/confound removal approaches were compared using a non-parametric Friedman ANOVA for each of the assessed parameters (correlations and absolute differences, each assessed at subject and connection level (Supplementary figures S5–S7), as well as residual variance in the time series). To aggregate these findings, the individual approaches were ranked according to their
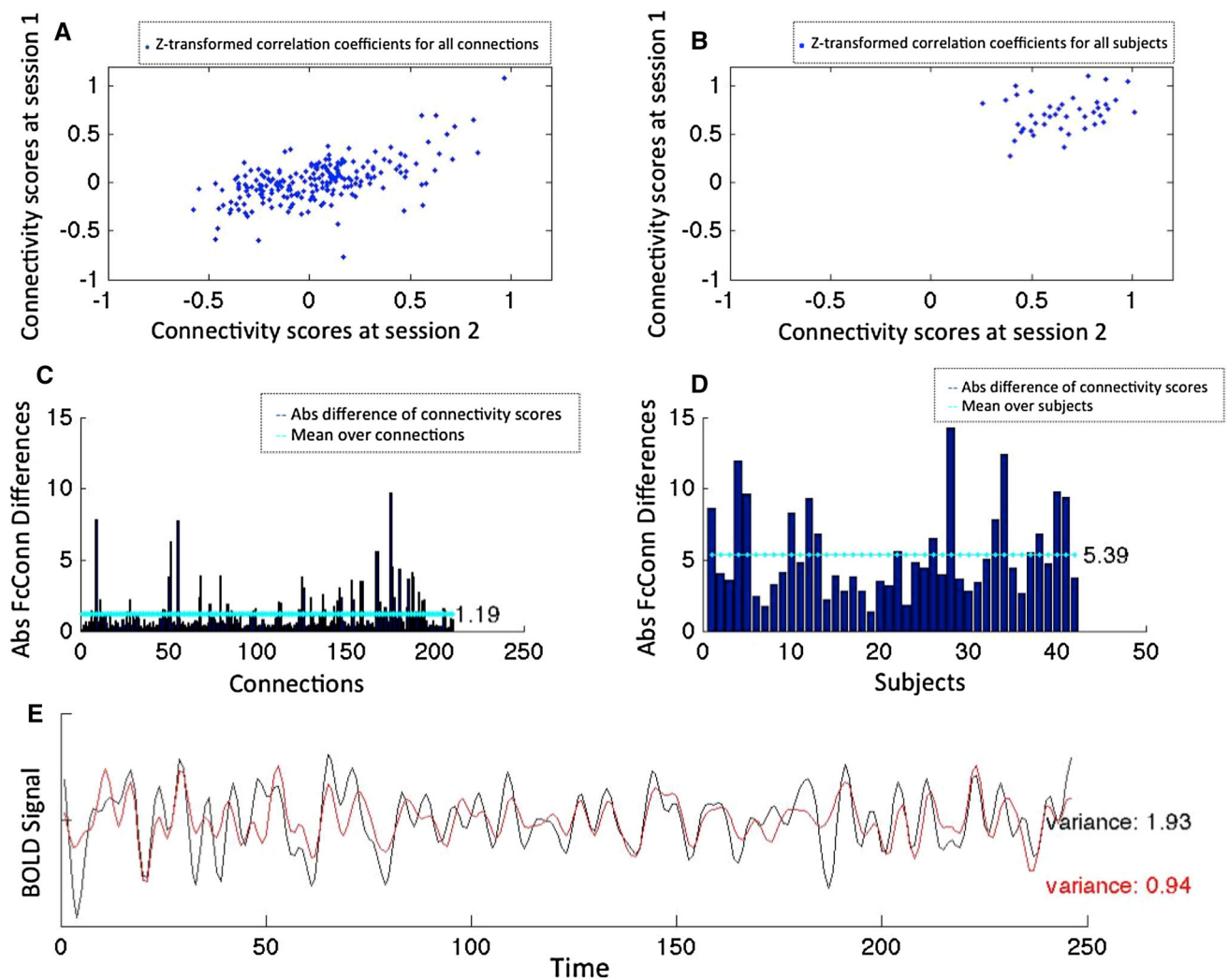
Fig. 5 Indices of the reliability: the four indices of reliability used here are shown. **a**, **b** Functional connectivity at two time points (**a**) at connection level, i.e., for all connections within a given subject (**b**) at subject level, i.e., for all the subjects within a given connection [here between left and right anterior insula (LaIns–RaIns)]. **c**, **d** Absolute differences of functional connectivity scores between the two sessions (**c**) at the connection level, i.e., the mean of the absolute differences over subjects for the 210 connections, and **d** at the subject level, i.e., the mean of the absolute differences over connections for the 42 subjects. **e** The variance within the BOLD signal time series of the left anterior Insula for two different combinations of signal processing methods ["GrpGM NoPCA NoGSR" (*black*), "NoGM PCA TSR" (*red*)]

reliability scores for each parameter. Subsequently, these reliability ranks were added over the different perspectives to obtain an overall reliability ranking. The overall reliability ranks allowed to identify reliable combination of different confound removal approaches at different perspectives.

Supplementary analysis

Given that the focus of our study was to investigate, which (combination of) analytical choices result in the best test–retest reliability functional connectivity estimates for meta-analytically defined networks, the main analyses used the entire significant clusters of the previously defined eSAD

and WM networks as regions of interest (ROIs). Acknowledging the alternative strategy of representing these ROIs by spheres around their center coordinates, we then repeated all analyses using spherical ROIs of 5 mm radius.

Results

The setup of our study allows us to perform a large number of different analyses. We first provide an overview on the test–retest reliability as reflected by the two different measures, i.e., rank-correlations and absolute differences. Here, the rankings based on the reliability of subject order

(RoSO) and those based on the reliability of connection order (RoCO) are combined. Next, we present an overview on the reliability from either perspective, combining the two measures. Finally, we provide the overall summary together with the ranking based on the residual variance in the time series as well as the information on the proportion of positive vs. negative connections. The individual test retest rankings by the two different methods and different perspectives are presented in the supplementary results (cf. Supplementary figures S1–S4).

In addition, we would like to note that we present findings for "within-network" and "between-network" connectivity. The former represents a summary of the rankings obtained for the extended socio-affective default mode as well as the working memory network, each showing strong, positive coupling among their respective nodes. The latter represents the connections between all possible pairs of nodes from either of these two major networks that are often conceptualized as being antagonistic to each other.

Reliability using different indices

The combined ranks, based on Kendall's rank correlations as the measure of subject- and connection-order, are shown in Fig. 6. The approaches are ordered such that the most reliable method is placed on the top, the least reliable on the bottom. It may be noted that for both within- and between-network connections, PCA denoising seems to have a rather detrimental effect on test–retest reliability, as most combinations including PCA denoising rank in the lower half and none is found in the top 10. On the other hand, gray matter masking, which is part of more than half of the ten most reliable approaches, seems to improve reliability. In particular, individual gray matter masking for within-network connections and group gray matter masking for between-network connections provide a better reliability. Global signal removal seems to have detrimental effect on the overall pattern for both within- and between-network connections. No removal again provided the most reliable correlation values for between-network

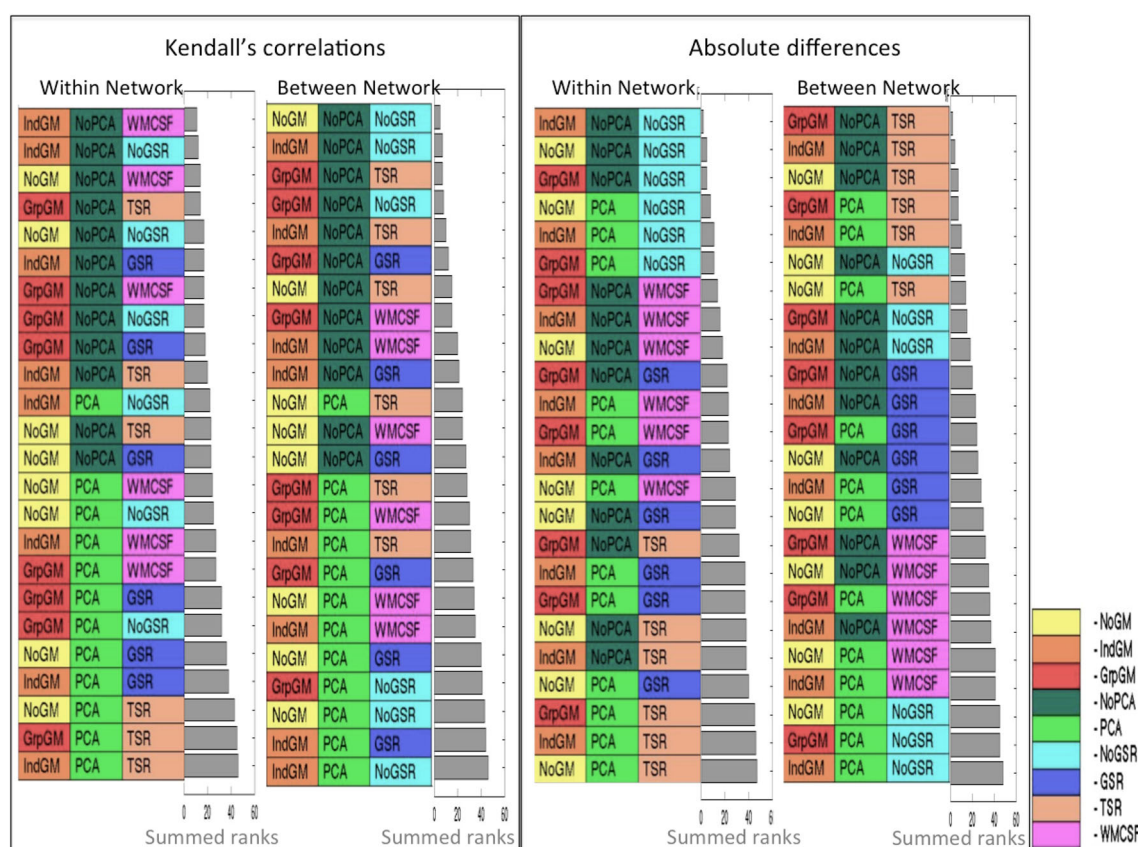

Fig. 6 Combined rankings of the test–retest reliability at the subject and connection level for Kendall's correlations and absolute differences. The "within networks" ranking refers to intra-network connections of the working memory and the default mode network and the "between networks" to inter-network connections. The *gray bar* represents the summed ranks for the respective categories

connections. Nevertheless, the rank-order stability of within-network connections was improved by removal of WM and CSF signals (WMCSF).

The assessment of reliability, by measuring absolute differences rather than measuring Kendall's correlations, corroborated most of these observations. In particular, we again found that using gray matter masking and refraining from PCA denoising yielded more reliable estimates of functional connectivity. While this pattern is not as clear-cut as for the correlation-based measure, it again held true for both within- and between-network connections. There is, however, a striking change in the overall pattern with respect to the effects of global signal removal. No removal again provided the most reliable absolute values for within-network connections. Nevertheless, the numerical stability of between-network connections was clearly improved by removing the global signal in all three-tissue classes (TSR).

Reliability from the subject and connection perspective

As noted in the methods, RoSO assesses how well the relative differentiation between subjects is reproduced at a second time point and is hence of particular relevance for between-subject analyses, e.g., in clinical application. In contrast, RoCO assesses how well the relative differentiation between connections in a particular subject is reproduced and is hence of particular relevance for within-subject analyses, e.g., in connectome modeling.

Several major trends of reliability noted in the previous section are again well observable in this analysis (Fig. 7). In particular, we again found that PCA denoising has a rather detrimental effect on reliability. In contrast, when considering within-network RoCO, PCA denoising has improved the reliability, namely, in the absence of global signal regression. Moreover, gray matter masking, in particular when using the mean tissue probabilities across the entire group, generally yields more reliable estimates of functional connectivity, although individual gray matter masking is more prominent when considering within-network connections, especially RoCO. With respect to the influence of global signal removal, we again found a more heterogeneous pattern with a clear distinction between within-network and between-network connections. With respect to the former, both RoSO and RoCO are highest when no global signal removal is performed, followed by approaches involving the removal of WM and CSF signals (WMCSF). For between-network connections, linear
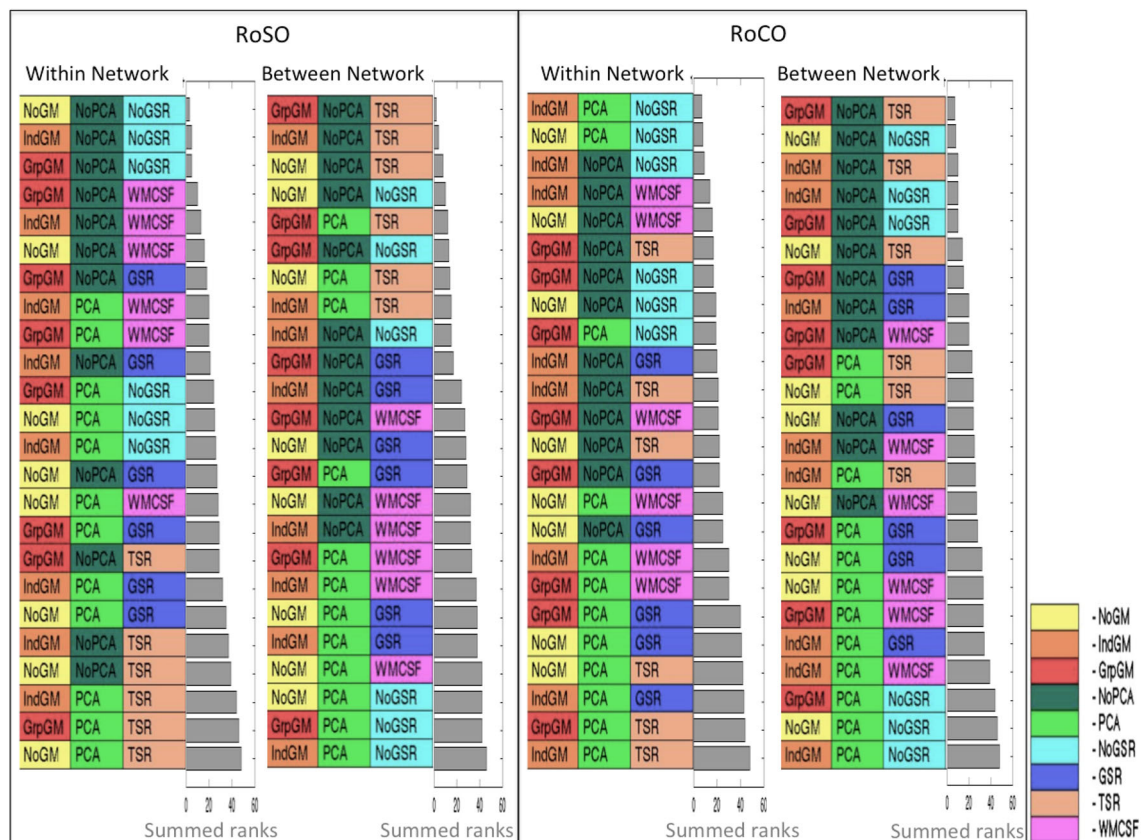

Fig. 7 Summary rankings for RoSO and RoCO. Reliability for within network (WMN and eSAD) and between networks is shown separately each combining Kendall's correlations and absolute difference. The *gray bar* represents the summed ranks for the respective categories

removal of the global signal for all three-tissue classes (TSR) yields the highest RoSO and RoCO, but for RoCO, neither removing any global signal nor performing a PCA denoising yields the highest reliability with no gray matter masking.

Summary of reliability ranking

The summary ranking across both indices (Kendall's correlations and absolute differences) and both perspectives (RoSO and RoCO) of reliability reflects the major patterns noted in the individual analyses (Fig. 8). Gray matter masking improves reliability. PCA denoising leads to lower test–retest reliability. Within-network connections are most reliably estimated when using no global signal regression and with removing the global WM and CSF signal representing the next-best approach. In contrast, between-network connections are most reliably measured by linear and second-order removal of global signals of all three-tissue classes.

Proportion of positive vs. negative connectivity scores and residual variance in the time series

Addressing the issue of anti-correlations, we assessed the proportion of positive vs. negative connections, i.e., connections with r (and hence Z-scores) below zero (Fig. 9). As expected, within-network connections are predominantly positive. It is, moreover, interesting to note that the least reliable approaches, i.e., those at the bottom of the list, also featured (somewhat) less consistent positive connections. The more striking observation, however, relates to the between-network connections. These are consistently negative when any form of global signal regression is used. If neither global signal regression nor PCA denoising are used, however, all connections are positive. Finally, when PCA denoising but no global signal regression is used, roughly half of the connections are positive.

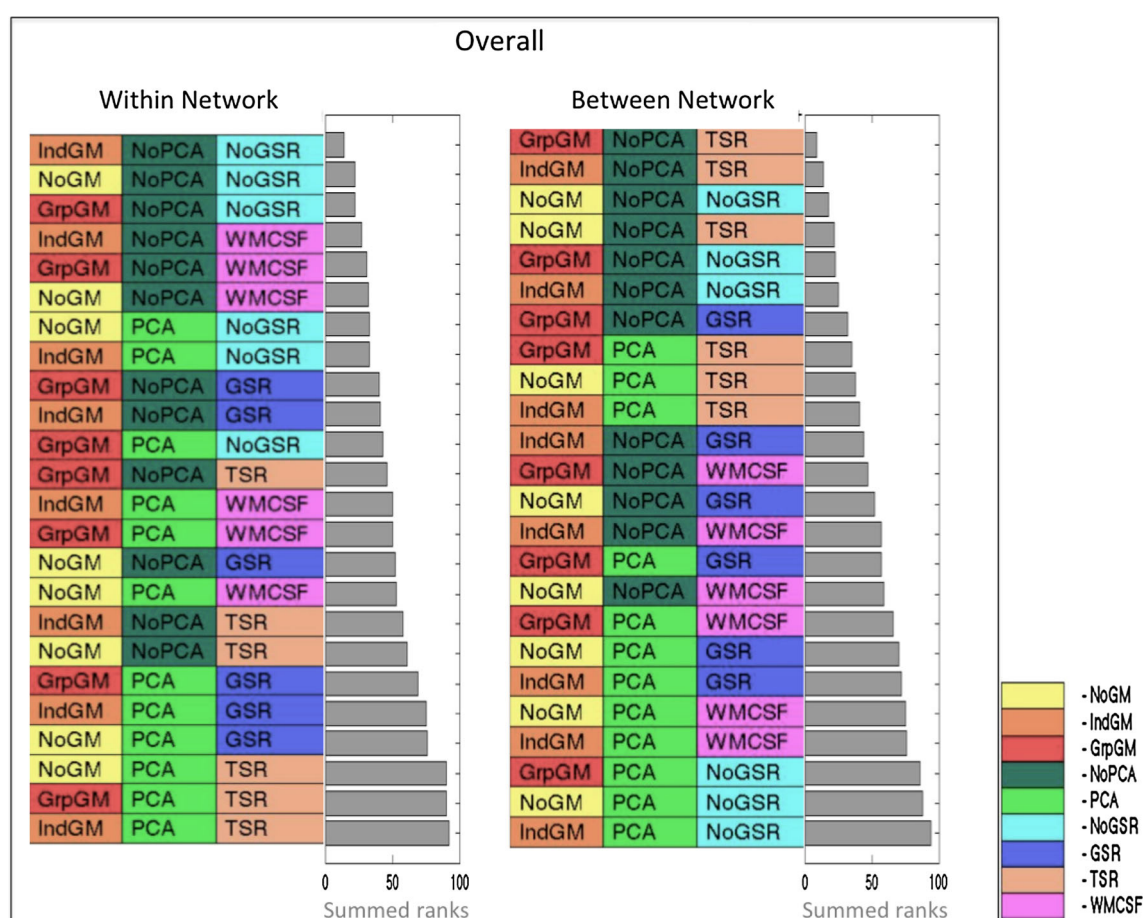Assessment of residual variance in the extracted time series expectedly reveals that refraining from PCA

Fig. 8 Summary rankings of reliability across Kendall's correlations and absolute differences as well as RoSO and RoCO, separately for within (WMN and eSAD) and between networks. The *gray bar* represents the summed ranks for the respective categories
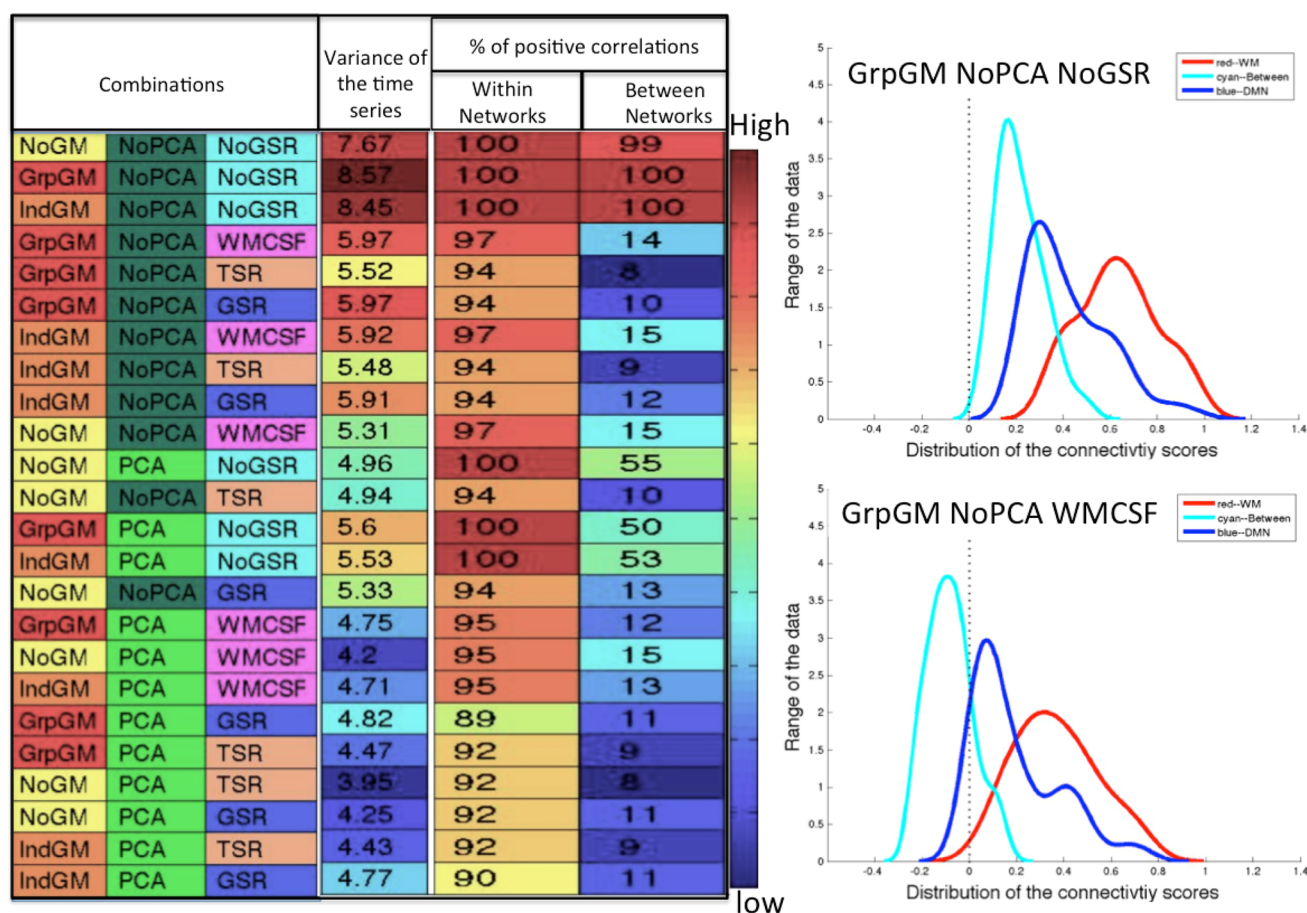
Combinations			Variance of the time series	% of positive correlations	
				Within Networks	Between Networks
NoGM	NoPCA	NoGSR	7.67	100	99
GrpGM	NoPCA	NoGSR	8.57	100	100
IndGM	NoPCA	NoGSR	8.45	100	100
GrpGM	NoPCA	WMCSF	5.97	97	14
GrpGM	NoPCA	TSR	5.52	94	8
GrpGM	NoPCA	GSR	5.97	94	10
IndGM	NoPCA	WMCSF	5.92	97	15
IndGM	NoPCA	TSR	5.48	94	9
IndGM	NoPCA	GSR	5.91	94	12
NoGM	NoPCA	WMCSF	5.31	97	15
NoGM	PCA	NoGSR	4.96	100	55
NoGM	NoPCA	TSR	4.94	94	10
GrpGM	PCA	NoGSR	5.6	100	50
IndGM	PCA	NoGSR	5.53	100	53
NoGM	NoPCA	GSR	5.33	94	13
GrpGM	PCA	WMCSF	4.75	95	12
NoGM	PCA	WMCSF	4.2	95	15
IndGM	PCA	WMCSF	4.71	95	13
GrpGM	PCA	GSR	4.82	89	11
GrpGM	PCA	TSR	4.47	92	9
NoGM	PCA	TSR	3.95	92	8
NoGM	PCA	GSR	4.25	92	11
IndGM	PCA	TSR	4.43	92	9
IndGM	PCA	GSR	4.77	90	11

Fig. 9 The variance left within the time series (far *left column*) and the percentage of positive correlations (*columns* on the far *right*) for both within and between networks arranged by the overall ranking of the reliability. The plots on the *right* side exemplify the difference of the distribution of the connectivity scores at different combinations ["GrpGM NoPCA NoGSR" (*top*), "GrpGM NoPCA WMCSF" (*bottom*)]

denoising and using no global signal regression retained more variance. Gray matter masking also seemed to perform well with regard to this measure.

Supplementary analysis

The results of the supplementary analysis conducted using spherical ROIs of 5 mm radius rather than the actual cluster volumes are detailed in the supplementary material. The summary ranking across both indices (Kendall's correlations and absolute differences) and both perspectives (RoSO and RoCO) of reliability reflect the major patterns noticed in the main analysis, except for the gray matter masking. The supplementary results associated with the PCA denoising and the mean global signal regression remain the same as in the main analysis. In turn, the supplementary results illustrate that using spherical ROI of 5 mm radius (i.e., smaller VOIs) favor No GM masking (cf. Supplementary figure S7).

Discussion

The key idea behind resting-state fMRI analyses is to estimate functional connectivity between distant brain regions based on the correlation of their BOLD time series (Biswal et al. 1995, 1997). The fundamental assumption behind this conceptualization is that the extracted time series reflect the effects of ongoing neuronal computation through hemodynamic coupling, such that correlated signal changes reflect inter-regional synchronization. However, systematic sources of non-neuronal fluctuations in EPI signals likely influence these functional connectivity estimates (Biswal et al. 1995; Friston et al. 1996; Fox and Raichle 2007; Buckner 2010; Cole et al. 2010). Addressing these non-neuronal signals is, therefore, a critical consideration in any functional connectivity approach. In this study, we investigated the influence of various preprocessing approaches meant to deal with this issue, including gray matter masking, PCA denoising, and global signal regression. Our findings are based on investigating two a

priori defined networks (the extended socio-affective default mode network and the working-memory network) in a sample of 42 subjects scanned twice, with an average retest-delay of 175 days. We found that gray matter masking based on group-average GM probabilities improved reliability, while confound removal approaches (either PCA denoising or global signal regression) reduced it. However, the study has yielded some mixed results that will be discussed in this section.

Recently, Shirer et al. (2015) investigated a confound removal pipeline that optimizes resting state fMRI data, which is comparable to our study. They performed a reliability study dealing with confound removal combined with various bandpass filter selections. In contrast, in this study, the focus is mainly on seed region time-series extraction methods based on different methods for gray-matter masking, combined with various confound removal techniques. There are several additional differences between both studies. Shirer et al. (2015) used ten components for the PCA model (5 from WM and 5 from CSF) and computed WM and CSF signals using a 3-mm radius spherical ROI centered on (arbitrary) WM and CSF regions. In contrast, we here used a five components PCA model, noting that five dominant principle components have been shown to effectively remove the relevant noise (Chai et al. 2012). Moreover, the mean WM and CSF signal was computed using the entire segmented WM and CSF regions in our study, assuming that signal from small regions may not model the appropriate noise term. In addition, they performed reliability analyses to evaluate the motion parameters, whereas we included them in the standard pre-processing given convincing previous evidence for using a 24-parameter motion regression model (Power et al. 2015; Satterthwaite et al. 2013) and bandpass filtered frequencies between 0.01 and 0.08 Hz (Biswal et al. 1995; Cordes et al. 2001; Fox et al. 2005; Zou et al. 2008; Van Dijk et al. 2012; Tsvetanov et al. 2015). Therefore, both studies deal with similar issues but address complementary aspects.

Different perspectives

Reliability of subjects (RoSO) and reliability of connections (RoCO) represent two fundamentally different views on reliability of resting-state measurements (Gorgolewski et al. 2013). Conceptually, assessing the RoSO allows us to identify which combinations of processing steps that yield a reproducible relationship between subjects for each connection, while RoCO identifies the combinations that yield the relationship between different connections in the same subject. RoSO is fundamental for any analysis focusing on between-subject differences. Example applications would include brain-phenotype associations, e.g.,

the correlation of connectivity estimates with neuropsychological or other behavioral measures (Müller et al. 2014), including clinical analyses comparing patients to healthy control subjects (Zhang and Raichle 2010; Hoptman et al. 2012; Müller et al. 2013). In contrast, RoCO is most relevant when performing any within-subject modeling, either as a primary goal, e.g., when performing connectivity-based parcellation, or to compute derivative measures characterizing the individual connectome (Eickhoff et al. 2011; Bzdok et al. 2013; Clos et al. 2013). Examples of the latter include graph-theory-based analyses that compute characteristic network measures from the individual connectome (Shen et al. 2010; Wang et al. 2011; Reid and Evans 2013). In other words, the results from the RoSO are particularly pertinent, when the focus is on group comparison or across-subject associations, whereas the results from the RoCO are relevant when the focus is on the structure of an individual subject's connectivity matrix.

Assessed (combinations of) signal processing steps

Here, we addressed the effects of gray-matter masking during the ROI time-series extraction (which has received rather little attention up to now), the influence of PCA denoising (which has at times been suggested but is not commonly used), and global signal regression (which is still highly controversial). The extracted ROI time series characterizes the temporal dynamics of the selected region as captured by the evoked BOLD response. While ROI time-series extraction plays a key role when studying the regional specific BOLD signal, the respective methods are rarely discussed even though it may affect reliability of subsequent analyses. For example, gray-matter masking is frequently used to restrict signal extraction to gray matter as much as possible, even though the benefits of doing so have not been explicitly demonstrated. In this study, we thus investigated this issue by examining the reliability of various gray-matter masking approaches.

Probably, the best-investigated source of spurious variance in RS time series is head motion (Van Dijk et al. 2012; Satterthwaite et al. 2013; Griffanti et al. 2014; Patriat et al. 2015; Power et al. 2015; Wong et al. 2016). Satterthwaite et al. (2013), using a 24-parameter motion regression approach, found that the first derivative as well as the quadratic effects of both realignment parameters and derivatives could account for these effects. In addition, in this study, the residual signal after removal of the variance associated with confounds variables is band pass filtered between 0.01 and 0.08 Hz, which is unfortunately known to be influenced by various noise components (Birn et al. 2006). Niazy et al. (2011) indicated that resting-state networks show temporal correlations across a wide frequency range, even though the resting-state networks are

dominated by low frequencies of the BOLD signal. However, there is ample evidence that the BOLD signal which is measured by fMRI and from which functional connectivity maps are derived is dominated by low-frequency fluctuations (Biswal et al. 1995; Cordes et al. 2001). Thus, to stay in line with standard applications, we followed the well-established standard of bandpass filtering and motion regression (Satterthwaite et al. 2013). Furthermore, it has been argued that global signal regression may be beneficial to deal with motion effects (Murphy et al. 2009; Power et al. 2012). In contrast, previous studies addressing the influence of global signal removal (Weissenbacher et al. 2009; Chai et al. 2012; Chen et al. 2012) and those assessing test–retest reliability (Shehzad et al. 2009; Gorgolewski et al. 2013; Birn et al. 2014) used less extensive motion regression protocols. Acknowledging new approaches based on automatically classifying and removing noise components have recently emerged (Behzadi et al. 2007), we here focused on three steps commonly used in settings in which physiological noise recording is not available and data quality is not sufficient for reliable estimation of noise components in individual subjects. Therefore, the paper aims to study the reliability and reproducibility of functional connectivity patterns in "clinical quality" data rather than in optimal datasets with low spatial and temporal resolution as well as physiological recordings.

Gray matter masking during time-series extraction

The time series extracted from an ROI represents the time-varying BOLD fluctuations within that region. Using one of the common approaches (Friston et al. 2006), we computed the first eigenvariate to obtain the characteristic time series for each ROI that accounts for the largest proportion of the variance in the set of voxel-wise time series. In general, voxels comprising the ROI may extend into the WM or CSF region, especially for a priori meta-analytically defined clusters, which usually do not respect the tissue class locations of the subjects under study. However, signals obtained from either WM or CSF voxels are not of interest in the functional connectivity analysis, as they should be of non-neuronal origin. One approach to reduce the influence of these unwanted signals and locally optimize the time-series extraction toward the biologically relevant voxels is to use gray matter masking. In that context, however, a fixed threshold for GM segmentation seems inappropriate, given that it could lead to exclusion of entire regions as well as having no effect in others. Our results indicates that using gray-matter masking when extracting the time series, i.e., considering only those voxels in the ROI that are above the median GM probability, yield more reliable connectivity scores.

Since there are no previous investigations into the effect of performing local optimization of ROIs toward gray-matter voxels, we here investigated two different approaches (median split based on the individual and group-averaged GM probabilities) and compared them to the "baseline" approach of using the entire ROI volume without masking. Factors like head motion could influence the outcome of various GM masks used for time-series extraction investigated in this study. Subjects with higher head motion may benefit either less (due to reduced fit) or more (due to poor individual segmentation) from the group-level GM masking. Therefore, the rationale for evaluating both approaches is that individual GM probabilities should best reflect a particular subject's anatomy after spatial normalization, but comes at the disadvantage of being potentially noisier given that they are based on a single scan. In contrast, group-level GM probabilities should be less specific but more robust. Our results are particularly true when the mean tissue probabilities across the entire group were used. In our view, this not only indicates the beneficial effects of gray matter masking and hence supports the aforementioned motivation to perform a local optimization, but also suggests that group-level masking, albeit potentially less specific, may be the preferable choice due to increased robustness. In addition, individual GM probabilities produce reliable results for within subject studies. Nevertheless, the segmentations and spatial normalization of the EPI images might be less precise as compared to that of high-resolution T1 images, due to the lower resolution and poorer contrast. This may entail somewhat higher registration inaccuracies, which, in turn, may have had some influence on the results. A straightforward and more traditional approach for gray matter masking would be to use a population-based a priori tissue mask (e.g., ICBM gray matter map). However, the use of such mask to define gray matter in the ROI may be more sensitive to (systematic) registration errors stemming, e.g., from differences in the studied population to the population that was used to construct the a priori tissue masks. In summary, we would thus recommend the use of a study specific group gray matter mask when dealing with large clusters such as derived from neuroimaging meta-analyses.

Interestingly, a somewhat different pattern emerges when representing the regions of interest not by the full highly threshold clusters derived from the meta-analyses (cf. Table 1) but rather by spheres of 5 mm radius around their peak coordinate (cf. Supplementary figures S5–S7). These definitions differ from those used in the main analysis in several aspects. In particular, these spherical ROIs contain a more uniform (compared to the cluster-based ones) and smaller number of voxels. In analyzing the effect of gray matter masking on these spherical ROIs, we found

that no masking yielded the best reliability and would propose two possible explanations (cf. Supplementary figures S5–S7). First, the smaller extent of these spherical ROIs most likely yielded a lower proportion of voxels located in WM and CSF, as indicated by a higher mean GM probability, although this is not a criterion for their definition. Second, given the smaller size of the spherical ROIs, the performed median split may have resulted in a critical further reduction of available voxels that renders the results unstable due to session-to-session misalignment, noise, or other factors. As a conclusion, it is advisable to implement gray matter masking for larger, a priori defined clusters based on the group-averaged GM probabilities to improve the reliability. In turn, when using smaller, spherical ROIs, no gray matter masking seems preferable.

PCA denoising

Cleaning the data with PCA denoising has been introduced by Behzadi et al. (2007) and frequently used since (e.g., Kellermann et al. 2013). In this study, we performed PCA denoising using the time course of the five most dominant principal components as confound regressors, effectively removing signal correlated with these. In an evaluation study, Chai et al. (2012) reported that removing principal components derived from WM and CSF regions is advisable to reduce the influence of physiologically induced artifacts, as components derived from WM and CSF regions are unlikely to include neural activity. In particular, it has been argued that physiologically induced artifacts should be particularly present within WM, ventricles, and large vessels (Chang et al. 2009). In addition, PCA denoising should remove effects that are widely distributed over the brain, including again variance related to physiological sources (Chai et al. 2012). Finally, it is worth mentioning that the first principle component is closely related to the global mean signal.

Our results focusing on test–retest reliability from two different perspectives (RoSO and RoCO), however, indicate that PCA denoising is not beneficial under either perspective, irrespectively of the remaining settings. These findings thus replicate the findings by Power et al. (2014) that PCA denoising does not yield encouraging results. In addition, Shirer et al. (2015) observed a decrease in test–retest reliability with PCA denoising. We note that, following the proposed method by Behzadi et al. (2007), the main analysis presented here obtain the principal components from the segmented white matter and CSF masks. As an alternative approach, principle components may also be computed from the whole brain mask, i.e., GM, WM, and CSF. We thus performed an additional analysis using PCA components derived from the entire brain, but observed similar results to those obtained from using WM/CSF

derived components (cf. Supplementary figures S8–S10). These results converge with those of Soltysik et al. (2015), which reveal that PCA extracted from whole brain yield similar results to those obtained from using WM and CSF regions. In summary, we would thus argue that PCA denoising has no beneficial effect on the test–retest reliability of RS-FC estimates, at least within the settings evaluated in this study. When investigating resting-state functional connectivity between a priori specified regions of interest refraining from PCA denoising should hence provide the more reliable results.

Global signal regression

Global signal regression, i.e., the removal of variance in the individual voxels' time series that can be explained by the average (global) signal across the entire brain, has become a controversial topic recently. Historically, it was based on the global scaling approaches utilized in the early (functional) PET studies, which were necessary to allow inference on localized and hence specific changes in blood flow. The key idea behind this approach has been retained in virtually all MRI-based neuroimaging studies, rendering global signal regression a common feature for both task- and resting-state fMRI. Similar to its origins in PET, the purpose is again to facilitate the detection of localized neuronal effects. Using GSR assumes that meaningful effects (reflecting activations or functional connectivity) are based on local variations in neuronal activity. Consequently, global signals, which are thought to mainly originate from physiological rather than neuronal sources, should be treated as a confounding influence. In line with this view, Power et al. (2014) observed that global signal regression is also an effective means of reducing motion-related effects in resting-state fMRI data.

Following the outlined logic, global signal removal has been the standard approach for many years until, more recently, it has been argued (Murphy et al. 2009; Weissenbacher et al. 2009; Saad et al. 2012) that GSR might introduce artificial anti-correlations. In addition, Chen et al. (2012) quantified the global noise levels, and based on the noise level within the data set, they advised to determine whether to include or exclude the global signal regressors based on this information. Ultimately, the issue of whether GSR should be employed or not remains contentious. Likewise, the effects of removing global vs. tissue-class specific mean signals, in particular only those for WM and CSF are still unclear. In this study, we thus investigated seven different variants of global signal removal involving global, mean tissue class and mean WM/CSF signal removal at the first or second order as well as no GSR.

Regarding the effects of global signal removal on test–retest reliability, our investigation yields somewhat mixed

results. Overall, we found that without any mean signal regression yields the highest reliability over both subjects and connections. However, when looking at the results in more detail, it may be noted that these overall findings are strongly driven by the within-network analyses. Here, not removing any GSR clearly yields the most reliable measures of functional connectivity. In turn, estimates for functional connectivity between the two assessed networks (WMN and eSAD) are most reliable when mean signal time courses for all three-tissue classes were removed from the data. Finally, we noted that removing the mean WM and CSF signal seems to provide a good compromise, as this approach yields reliable estimates of within- and between-network connections, although it is not the best approach in either case. Furthermore, Yan et al. (2013b) suggested that global signal regression is nearly identical to gray matter regression. Thus, both the results from Yan et al. (2013b) and our present data argue for using only the mean WM and CSF signal (but not the mean gray matter) for nuisance signal regression.

The issue of global signal regression is strongly tied to the question of (spurious or induced) anti-correlations. This is also evident in our data. Without any global signal removal, both within- and between-network connections correlate positively. This indicates that global fluctuations override any potential local anti-correlations. Yet, when variance explained by the global signal or the mean WM and CSF is removed, between-network connections become predominantly negative. That is, only when global changes in the BOLD signal are removed, do the estimated functional connectivity values reflect the repeatedly advocated anti-correlated structure of "task-positive" and "task-negative" networks. Should these thus be considered spurious? One argument against this rather critical view comes from task-based fMRI studies (Greicius et al. 2003; Greicius and Menon 2004), which have clearly shown that regions such as the eSAD reduce their activity during cognitive tasks, which in order recruit fronto-parietal networks such as the working-memory network investigated here. However, global signal removal or, more commonly, scaling is also a standard approach also in task-fMRI (Macey et al. 2004). Another possibility is that global signal may be comprised primarily of non-neuronal sources, rendering the positive correlation between any two parts of the brain in the absence of global signal regression spurious (Murphy et al. 2009). We would, therefore, argue that global (positive) correlation and between-network anti-correlations might be considered as two aspects of a more complex situation. In particular, it seems that anti-correlative structures between large-scale networks are superimposed on larger waves of global signal changes, which may be non-neuronal in origin (Fox et al. 2009). Nevertheless, more recently, Schölvinck et al. (2013)

suggested that the global signal is tightly coupled to the neuronal signal. In addition, Pisauro et al. 2016 showed that global components in mice are coupled to pupil dilation as a measure of sympathetic function. Thus, they may be partially neuronal and non-neuronal in origin. In such case, removal of global signals likewise acts as a focus on (smaller) local effects of anti-correlated nature while ignoring the large-scale synchronization of BOLD patterns. In turn, not removing any global signal would preserve the latter and hence bring the positive relation between all time-series that is present in the acquired data into focus.

General discussion

When assessing the test–retest reliability of resting-state fMRI connectivity estimates, one unlikely but still important caveat must be considered. It is possible that increased reliability, i.e., higher correlation and lower absolute difference, will be caused by excessive removal of variance. In the extreme case, when the time series would be reduced to a flat line, test–retest reliability would be perfect. However, beyond this hypothetical extreme case, the relationship between reliability and variance is interesting; as it sheds light on the question to what extent our methods remove noise (in that case residual variance and reliability would be positively related) or relevant signal (which would render the relationship negative). In our assessment, we found that methods providing results that are more reliable also feature higher residual variance within the extracted time series (Fig. 9, the correlation between residual variance and reliability scores is 0.87) Therefore, reliability seems proportional to the retained variance, reinforcing the observations by Birn et al. (2014) and Yan et al. (2013a).

Another point to consider is the relationship between reliability and validity. The underlying idea of all preprocessing approaches is to remove variance in the data that may be attributable to noise or, more generally, non-neuronal sources. This naively assumes that more aggressive confound removal should increase the biological validity of the obtained results. However, this assumption has been challenged, most notably with respect to global signal regression. Here, it has been argued that removing global signal as a confound may actually introduce a bias in the analysis (Murphy et al. 2009; Weissenbacher et al. 2009; Saad et al. 2012), that may lead to reduction in validity. Conversely, the argument has been made that global signal regression is the most effective approach to remove the effect of motion-related variance (Power et al. 2014) and hence should increase validity. This already illustrates that the relationship between data preprocessing, and in particular confound removal, and validity is not trivial. The present results add another layer of complexity by showing

that refraining from using global signal regression and PCA denoising, i.e., using less confounds removal, actually lead to better test–retest reliability. In other words, removing variance that is related to potentially confounding factors reduces reliability, pointing to the possibility that structured noise may be beneficial for test–retest reliability. In addition, indeed, it may be assumed that vascular of physiological factors remain largely stable between sessions and hence help to increase reliability, even though their removal should, in theory, improve the validity of the results. Maximizing (test–retest) reliability and biological specificity/validity may hence represent (partially) conflicting aims.

The functional connectivity strength (i.e., correlation coefficients) between regions might vary with changes in the level of observation noise (Friston 2011). In this study, two resting-state networks (eSAD and WMN), which may be considered as robustly a priori defined resting state networks has been chosen, with prior assumptions such as strong positive coupling among them and anti-correlated with each other (Fox et al. 2005). When there is not any change in the observational noise, then the functional connectivity strength (i.e., correlation coefficients) is expected to be stable (Friston 2011). Therefore, instead of quantifying the connectivity strengths, we mainly focused on reproducibility of the connectivity strength with a certain confound removal within a subject from one session to another session. Furthermore, following the current standard in the field, our study quantified functional connectivity by the Pearson correlations between the time series of two regions. Consequently, other regions within or outside the network could influence such correlations. Such influences, however, were not specifically investigated, given that they should be likewise present in both sessions and, most importantly, the focus of our work is to provide an assessment of how the reproducibility of the widely used time-series correlation measures are based on different approaches to confound removal. That is, we here addressed the pragmatic question, which confound removal strategy yields the highest reliability for a standard analysis approach, rather than addressing which analysis approach may yield the most appropriate representation of a network. Evidently, more investigations are needed to better understand the sources of both noise and signal in resting-state fMRI data, a question that is complicated by a lack of ground truth. Nevertheless, the current results thus point to a potential tradeoff between reliability (which may benefit from structured noise) and biological validity (which should be optimal if all non-neuronal variance is removed (Huettel et al. 2004; Chang et al. 2009; Kim and Ogawa 2012). Based on the present results, we would thus tentatively propose that in cases in which reliability should be of particular importance, for example, in clinical applications,

it may be advisable to refrain from global signal regression and PCA denoising to maximize the reliability albeit potentially through the influence of structured noise.

ICA-based denoising is one of the recently emerging confound removal approaches. A recent study showed that it can effectively remove the artifacts coupled with motion (Pruim et al. 2015b) and potential other sources of noise (Griffanti et al. 2014). The entire resting-state scan is decomposed into independent components (IC) (using FSL melodic, http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/MELODIC). ICs coupled with various artifacts were identified with the help of a classifier. ICs classified as noise is then regressed out of the raw fMRI time series. Thus, ICA-based denoising aims to automatically classify and remove the components representing mostly noise rather than neuronal signal (Salimi-Khorshidi et al. 2014; Pruim et al. 2015b). The effectiveness of the strategies mainly depends on the feature selection and the sensitivity of the classifier, as these parameters play a major role in identifying the artifactual signals. In recent evaluation studies, ICA-based denoising strategies resulted in an increase of the between subjects reproducibility (Griffanti et al. 2014; Pruim et al. 2015a). In this study, however, we did not address ICA-based denoising approaches, as we mainly focused on the currently most widely used approaches. In turn, ICA-based denoising is a very promising but yet emerging approach as also demonstrated in our survey. Therefore, further investigations are needed to address the reliability of ICA-based approach both in comparison to and in combination with conventional confound removal strategies. Along with it, there are methods that mainly address local and global artifacts induced by the hardware and partial volume effects (such as: ANATICOR (Jo et al. 2010, 2013)). As the current study mainly studied the influences of biologically induced artifacts, methods like ANATICOR were not addressed here. Furthermore, it has been observed from the literature survey (Fig. 1) that ANATICOR (which has been reported in the categories named 'others') is not a standard method and poorly used in the recent studies.

In this study, the connectivity measures were obtained with standard Pearson correlations. Other approaches have also been applied to this computation, with partial correlation becoming an increasingly advocated alternative (Cole et al. 2010). Partial correlation computes the correlations between two ROIs after regressing out the shared variance of all other ROI time series in the model. However, we are here concerned with testing the effects of several widely used analysis-choices on the reliability of the most common approach. Therefore, given that the overwhelming majority of all resting-state analyses employ full correlations, we here performed a practical evaluation of the impact of currently debated analyses choices on the estimation of functional connectivity by Pearson

correlations. Nevertheless, testing the test–retest reliability using partial correlation could be one perspective study of the current one. Furthermore, the subjects were instructed to close their eyes during the resting state session, to reduce the external (visual) stimulation and eye movements. All the subjects included in this study had confirmed to be awake while debriefing. The condition of eyes closed (EC) may be considered as a limitation of the study, as Patriat et al. (2013) showed higher reliability with eyes open (EO) condition rather than eyes closed (EC) condition. However, Patriat et al. (2013) also reported that the connectivity strengths are not sensitive to the global noise variations. Therefore, further investigations of reliability of EO and EC with and without global noise regression are needed to provide recommendations regarding this parameter. Finally, it has to be noted that the recommendations in this paper may not necessarily apply to brain-behavior analysis examining the relationship between behavioral measures and functional connectivity measures. That is, we here focused on a priori defined meta-analytical networks and their (known) relationships to each other as large-scale anti-correlated systems in the human brain (Fox et al. 2005). What remains to be assessed using a dedicated sample for which test–retest data not only of imaging measures, but also behavioral information is available is this, whether the methods yielding the best reliability in our analysis also provide the most reliable brain-behavior relationships. Likewise, it remains to be tested, whether the identified recommendations also hold for multivariate analyses, e.g., in the context of group classification.

Conclusions

This study assessed test–retest reliability of resting-state fMRI analyses based on a priori ROIs using methods that are applicable without direct recordings of physiological signals (heartbeat, breathing), as is common in clinical and neuro-scientific practice. In particular, our results showed that, when using the larger clusters as regions of interest, gray matter masking based on the group-average GM probabilities is advisable. However, In addition, PCA denoising reduces the reliability of connectivity estimates. Finally, with respect to global signal regression, we observed that refraining from this approach enhances test–retest reliability but comes at the expense of potentially poorer biological validity, including missing anti-correlations between what has been previously described as antagonistic networks. Here, removal of global white matter and CSF signals seems to provide a good compromise, as this approach yielded more reliable and potentially meaningful estimates of within- and between-network connections. Importantly, we note that reliability is

proportional to the retained variance, presumably including structured noise. Consequently, a compromise exists between maximizing the test–retest reliability and removing variance that may be attributable to non-neuronal sources.

Compliance with ethical standards

References

Amft M, Bzdok D, Laird AR et al (2015) Definition and characterization of an extended social-affective default network. Brain Struct Funct 220:1031–1049. doi:10.1007/s00429-013-0698-0

Ashburner J, Friston KJ (2005) Unified segmentation. Neuroimage 26:839–851. doi:10.1016/j.neuroimage.2005.02.018

Barch DM, Burgess GC, Harms MP et al (2013) Function in the human connectome: task-fMRI and individual differences in behavior. Neuroimage 80:169–189. doi:10.1016/j.neuroimage.2013.05.033

Behzadi Y, Restom K, Liau J, Liu TT (2007) A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. Neuroimage 37:90–101. doi:10.1016/j.neuroimage.2007.04.042

Birn RM, Diamond JB, Smith MABP (2006) Separating respiratory-variation-related fluctuations from neuronal-activity-related fluctuations in fMRI. Neuroimage 31:1536–1548

Birn RM, Molloy EK, Patriat R et al (2013) The effect of scan length on the reliability of resting-state fMRI connectivity estimates. Neuroimage 83:550–558. doi:10.1016/j.neuroimage.2013.05.099

Birn RM, Cornejo MD, Molloy EK et al (2014) The influence of physiological noise correction on test–retest reliability of resting-state functional connectivity. Brain Connect 4:511–522. doi:10.1089/brain.2014.0284

Biswal B, Yetkin FZ, Haughton VM, Hyde JS (1995) Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. Magn Reson Med 34:537–541. doi:10.1002/mrm.1910340409

Biswal BB, Van Kylen J, Hyde JS (1997) Simultaneous assessment of flow and BOLD signals in resting-state functional connectivity maps. NMR Biomed 10(4–5):165–170

Buckner RL (2010) Human functional connectivity: new tools, unresolved questions. Proc Natl Acad Sci 107(24):10769–10770

Buckner RL, Bandettini PA, O'Craven KM et al (1996) Detection of cortical activation during averaged single trials of a cognitive

task using functional magnetic resonance imaging. Proc Natl Acad Sci USA 93:14878–14883. doi:10.1073/pnas.93.25.14878

Button KS, Ioannidis JPA, Mokrysz C et al (2013a) Power failure: why small sample size undermines the reliability of neuroscience. Nat Rev Neurosci 14:365–376. doi:10.1038/nrn3475

Button KS, Ioannidis JPA, Mokrysz C et al (2013b) Empirical evidence for low reproducibility indicates low pre-study odds. Nat Rev Neurosci 14:877. doi:10.1038/nrn3475-c6

Bzdok D, Laird AR, Zilles K et al (2013) An investigation of the structural, connectional, and functional subspecialization in the human amygdala. Hum Brain Mapp 34:3247–3266. doi:10.1002/hbm.22138

Chai XJ, Castañón AN, Öngür D, Whitfield-Gabrieli S (2012) Anticorrelations in resting state networks without global signal regression. Neuroimage 59:1420–1428. doi:10.1016/j.neuroimage.2011.08.048

Chang C, Cunningham JP, Glover GH (2009) Influence of heart rate on the BOLD signal: the cardiac response function. Neuroimage 44:857–869. doi:10.1016/j.neuroimage.2008.09.029

Chen G, Chen G, Xie C et al (2012) A method to determine the necessity for global signal regression in resting-state fMRI studies. Magn Reson Med 68:1828–1835. doi:10.1002/mrm.24201

Chen B, Xu T, Zhou C et al (2015) Individual variability and test–retest reliability revealed by ten repeated resting-state brain scans over one month. PLoS One 10:e0144963. doi:10.1371/journal.pone.0144963

Clos M, Amunts K, Laird AR et al (2013) Tackling the multifunctional nature of Broca's region meta-analytically: co-activation-based parcellation of area 44. Neuroimage 83:174–188. doi:10.1016/j.neuroimage.2013.06.041

Cole DM, Smith SM, Beckmann CF (2010) Advances and pitfalls in the analysis and interpretation of resting-state FMRI data. Front Syst Neurosci 4:8. doi:10.3389/fnsys.2010.00008

Cordes D, Haughton VM, Arfanakis K et al (2001) Frequencies contributing to functional connectivity in the cerebral cortex in HResting-state data. Am J Neuroradiol 22:1326–1333

Eickhoff SB, Stephan KE, Mohlberg H et al (2005) A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. Neuroimage 25:1325–1335. doi:10.1016/j.neuroimage.2004.12.034

Eickhoff SB, Laird AR, Grefkes C et al (2009) Coordinate-based activation likelihood estimation meta-analysis of neuroimaging data: a random-effects approach based on empirical estimates of spatial uncertainty. Hum Brain Mapp 30:2907–2926. doi:10.1002/hbm.20718

Eickhoff SB, Bzdok D, Laird AR et al (2011) Co-activation patterns distinguish cortical modules, their connectivity and functional differentiation. Neuroimage 57:938–949. doi:10.1016/j.neuroimage.2011.05.021

Eickhoff SB, Bzdok D, Laird AR et al (2012) Activation likelihood estimation meta-analysis revisited. Neuroimage 59:2349–2361. doi:10.1016/j.neuroimage.2011.09.017

Eickhoff SB, Thirion B, Varoquaux G, Bzdok D (2015) Connectivity-based parcellation: critique and implications. Hum Brain Mapp 36:4771–4792

Fox MD, Raichle ME (2007) Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. Nat Rev Neurosci 8:700–711. doi:10.1038/nrn2201

Fox MD, Snyder AZ, Vincent JL et al (2005) The human brain is intrinsically organized into dynamic, anticorrelated functional networks. Proc Natl Acad Sci USA 102:9673–9678. doi:10.1073/pnas.0504136102

Fox MD, Zhang D, Snyder AZ, Raichle ME (2009) The global signal and observed anticorrelated resting state brain networks. J Neurophysiol 101:3270–3283. doi:10.1152/jn.90777.2008

Fox MD, Zhang D, Snyder AZ, Raichle ME (2013) The global signal and observed anticorrelated resting state brain networks 101:3270–3283. doi:10.1152/jn.90777.2008

Fox PT, Lancaster JL, Laird AR, Eickhoff SB (2014) Meta-analysis in human neuroimaging: computational modeling of large-scale databases. Annu Rev Neurosci 37:409–434. doi:10.1146/annurev-neuro-062012-170320

Frässle S, Paulus FM, Krach S, Jansen A (2016) Test–retest reliability of effective connectivity in the face perception network. Hum Brain Mapp 744:730–744. doi:10.1002/hbm.23061

Friston KJ (2011) Functional and effective connectivity: a review. Brain Connect 1:13–36. doi:10.1089/brain.2011.0008

Friston KJ, Williams S, Howard R et al (1996) Movement-related effects in fMRI time-series. Magn Reson Med 35:346–355. doi:10.1002/mrm.1910350312

Friston KJ, Rotshtein P, Geng JJ et al (2006) A critique of functional localisers. Neuroimage 30:1077–1087. doi:10.1016/j.neuroimage.2005.08.012

Gorgolewski KJ, Storkey AJ, Bastin ME et al (2013) Single subject fMRI test–retest reliability metrics and confounding factors. Neuroimage 69:231–243. doi:10.1016/j.neuroimage.2012.10.085

Greicius MD, Menon V (2004) Default-mode activity during a passive sensory task: uncoupled from deactivation but impacting activation. J Cogn Neurosci 16:1484–1492. doi:10.1162/0898929042568532

Greicius MD, Krasnow B, Reiss AL, Menon V (2003) Functional connectivity in the resting brain: a network analysis of the default mode hypothesis. Proc Natl Acad Sci USA 100:253–258. doi:10.1073/pnas.0135058100

Griffanti L, Salimi-Khorshidi G, Beckmann CF et al (2014) ICA-based artefact removal and accelerated fMRI acquisition for improved resting state network imaging. Neuroimage 95:232–247. doi:10.1016/j.neuroimage.2014.03.034

Guo WB, Liu F, Xue ZM et al (2011) Abnormal neural activities in first-episode, treatment-naïve, short-illness-duration, and treatment-response patients with major depressive disorder: a resting-state fMRI study. J Affect Disord 135:326–331. doi:10.1016/j.jad.2011.06.048

He B, Wang Y, Wu D (1999) Estimating cortical potentials from scalp EEG's in a realistically shaped inhomogeneous head model by means of the boundary element method. IEEE Trans Biomed Eng 46:1264–1268. doi:10.1109/10.790505

Hinke RM, Hu X, Stillman AE et al (1993) Functional magnetic resonance imaging of Broca's area during internal speech. NeuroReport 4:675–678

Hoptman MJ, Zuo X-N, D'Angelo D et al (2012) Decreased interhemispheric coordination in schizophrenia: a resting state fMRI study. Schizophr Res 141:1–7. doi:10.1016/j.schres.2012.07.027

Huettel SA, Song AW, McCarthy G (2004) Functional magnetic resonance imaging. Sinauer Associates Publishers, Sunderland

Jo HJ, Saad ZS, Simmons WK et al (2010) Mapping sources of correlation in resting state FMRI, with artifact detection and removal. Neuroimage 52:571–582. doi:10.1016/j.neuroimage.2010.04.246

Jo HJ, Gotts SJ, Reynolds RC et al (2013) Effective preprocessing procedures virtually eliminate distance-dependent motion artifacts in resting state FMRI. J Appl Math. doi:10.1155/2013/935154

Kellermann TS, Caspers S, Fox PT et al (2013) Task- and resting-state functional connectivity of brain regions related to affection and susceptible to concurrent cognitive demand. Neuroimage 72:69–82. doi:10.1016/j.neuroimage.2013.01.046

Kim S-G, Ogawa S (2012) Biophysical and physiological origins of blood oxygenation level-dependent fMRI signals. J Cereb Blood Flow Metab 32:1188–1206. doi:10.1038/jcbfm.2012.23

Kwong KK, Belliveau JW, Chesler DA et al (1992) Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation. Proc Natl Acad Sci USA 89:5675–5679. doi:10.1073/pnas.89.12.5675

Lai Y, Zhang X, van Drongelen W et al (2011) Noninvasive cortical imaging of epileptiform activities from interictal spikes in pediatric patients. Neuroimage 54:244–252. doi:10.1016/j.neuroimage.2010.07.026

Li Z, Kadivar A, Pluta J et al (2012) Test–retest stability analysis of resting brain activity revealed by blood oxygen level-dependent functional MRI. J Magn Reson Imaging 36:344–354. doi:10.1002/jmri.23670

Lu Y, Worrell GA, Zhang HC et al (2014) Noninvasive imaging of the high frequency brain activity in focal epilepsy patients. IEEE Trans Biomed Eng 61:1660–1667. doi:10.1109/TBME.2013.2297332

Macey PM, Macey KE, Kumar R, Harper RM (2004) A method for removal of global effects from fMRI time series. Neuroimage 22:360–366. doi:10.1016/j.neuroimage.2003.12.042

Müller VI, Cieslik EC, Laird AR et al (2013) Dysregulated left inferior parietal activity in schizophrenia and depression: functional connectivity and characterization. Front Hum Neurosci 7:268. doi:10.3389/fnhum.2013.00268

Müller VI, Langner R, Cieslik EC et al (2014) Interindividual differences in cognitive flexibility: influence of gray matter volume, functional connectivity and trait impulsivity. Brain Struct Funct 220:2401–2414. doi:10.1007/s00429-014-0797-6

Murphy K, Birn RM, Handwerker DA et al (2009) The impact of global signal regression on resting state correlations: are anti-correlated networks introduced? Neuroimage 44:893–905. doi:10.1016/j.neuroimage.2008.09.036

Niazy RK, Xie J, Miller K et al (2011) Spectral characteristics of resting state networks, 1st edn. Elsevier B.V, Amsterdam

Patriat R, Molloy EK, Meier TB et al (2013) The effect of resting condition on resting-state fMRI reliability and consistency: a comparison between resting with eyes open, closed, and fixated. Neuroimage 78:463–473. doi:10.1016/j.neuroimage.2013.04.013

Patriat R, Molloy EK, Birn R (2015) Using edge voxel information to improve motion regression for rs-fMRI connectivity studies. Brain Connect 5:582–595. doi:10.1089/brain.2014.0321

Pisauro MA, Benucci A, Carandini M (2016) Local and global contributions to hemodynamic activity in mouse cortex. J Neurophysiol 115:2931–2936. doi:10.1152/jn.00125.2016

Power JD, Barnes KA, Snyder AZ, Schlaggar BL, Petersen SE (2012) Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. Neuroimage 59:2142–2154. doi:10.1016/j.neuroimage.2011.10.018.Spurious

Power JD, Mitra A, Laumann TO et al (2014) Methods to detect, characterize, and remove motion artifact in resting state fMRI. Neuroimage 84:320–341. doi:10.1016/j.neuroimage.2013.08.048

Power JD, Schlaggar BL, Petersen SE (2015) Recent progress and outstanding issues in motion correction in resting state fMRI. Neuroimage 105:536–551. doi:10.1016/j.neuroimage.2014.10.044

Pruim RHR, Mennes M, Buitelaar JK, Beckmann CF (2015a) Evaluation of ICA-AROMA and alternative strategies for motion artifact removal in resting state fMRI. Neuroimage 112:278–287. doi:10.1016/j.neuroimage.2015.02.063

Pruim RHR, Mennes M, van Rooij D et al (2015b) ICA-AROMA: a robust ICA-based strategy for removing motion artifacts from fMRI data. Neuroimage 112:267–277. doi:10.1016/j.neuroimage.2015.02.064

Raemaekers M, Du Plessis S, Ramsey NF et al (2012) Test–retest variability underlying fMRI measurements. Neuroimage 60:717–727. doi:10.1016/j.neuroimage.2011.11.061

Reid AT, Evans AC (2013) Structural networks in Alzheimer's disease. Eur Neuropsychopharmacol 23:63–77. doi:10.1016/j.euroneuro.2012.11.010

Reid AT, Bzdok D, Langner R et al (2016) Multimodal connectivity mapping of the human left anterior and posterior lateral prefrontal cortex. Brain Struct Funct 221:2589–2605

Rottschy C, Langner R, Dogan I et al (2012) Modelling neural correlates of working memory: a coordinate-based meta-analysis. Neuroimage 60:830–846. doi:10.1016/j.neuroimage.2011.11.050

Saad ZS, Gotts SJ, Murphy K et al (2012) Trouble at rest: how correlation patterns and group differences become distorted after global signal regression. Brain Connect 2:25–32. doi:10.1089/brain.2012.0080

Salimi-Khorshidi G, Douaud G, Beckmann CF et al (2014) Automatic denoising of functional MRI data: combining independent component analysis and hierarchical fusion of classifiers. Neuroimage 90:449–468. doi:10.1016/j.neuroimage.2013.11.046

Satterthwaite TD, Elliott MA, Gerraty RT, Ruparel K, Loughead J, Calkins ME, Eickhoff SB, Hakonarson H, Gut RC, Gur REWD (2013) An improved framework for confound regression and filtering for control of motion artifact in the preprocessing of resting-state functional connectivity data. Neuroimage 64:240–256

Schilbach L (2016) Differential patterns of dysconnectivity in mirror neuron and mentalizing networks in schizophrenia. Schizophr Bull. doi:10.1093/schbul/sbw015

Schilbach L, Bzdok D, Timmermans B et al (2012) Introspective minds: using ALE meta-analyses to study commonalities in the neural correlates of emotional processing, social and unconstrained cognition. PLoS One 7:e30920. doi:10.1371/journal.pone.0030920

Schilbach L, Müller VI, Hoffstaedter F et al (2014) Meta-analytically informed network analysis of resting state fMRI reveals hyperconnectivity in an introspective socio-affective network in depression. PLoS One 9:e94973. doi:10.1371/journal.pone.0094973

Schölvinck ML, Leopold DA, Brookes MJ, Khader PH (2013) The contribution of electrophysiology to functional connectivity mapping. Neuroimage 80:297–306. doi:10.1016/j.neuroimage.2013.04.010

Shehzad Z, Kelly AMC, Reiss PT et al (2009) The resting brain: unconstrained yet reliable. Cereb Cortex 19:2209–2229. doi:10.1093/cercor/bhn256

Shen X, Papademetris X, Constable RT (2010) Graph-theory based parcellation of functional subunits in the brain from resting-state fMRI data. Neuroimage 50:1027–1035. doi:10.1016/j.neuroimage.2009.12.119

Shirer WR, Jiang H, Price CM et al (2015) Optimization of rs-fMRI pre-processing for enhanced signal-noise separation, test–retest reliability, and group discrimination. Neuroimage 117:67–79. doi:10.1016/j.neuroimage.2015.05.015

Smith DV, Utevsky AV, Bland AR et al (2014) Characterizing individual differences in functional connectivity using dual-regression and seed-based approaches. Neuroimage 95:1–12. doi:10.1016/j.neuroimage.2014.03.042

Soltysik DA, Thomasson D, Rajan S, Biassou N (2015) Improving the use of principal component analysis to reduce physiological noise and motion artifacts to increase the sensitivity of task-based fMRI. J Neurosci Methods 241:18–29. doi:10.1016/j.jneumeth.2014.11.015

Song X, Lawrence PP, Chen N-K (2016) Data-driven and predefined ROI-based quantification of long-term resting-state fMRI reproducibility. Brain Connect XX:1–16. doi:10.1089/brain.2015.0349

Thomason ME, Dennis EL, Joshi AA et al (2011) Resting-state fMRI can reliably map neural networks in children. Neuroimage 55:165–175. doi:10.1016/j.neuroimage.2010.11.080

Tsvetanov KA, Henson RNA, Tyler LK et al (2015) The effect of ageing on fMRI: correction for the confounding effects of vascular reactivity evaluated by joint fMRI and MEG in 335 adults. Hum Brain Mapp 36:2248–2269. doi:10.1002/hbm.22768

Van Dijk KR, Sabuncu MR, Buckner RL (2012) The influence of head motion on intrinsic functional connectivity MRI. Neuroimage 59:431–438. doi:10.1016/j.neuroimage.2011.07.044.The

Wang J-H, Zuo X-N, Gohel S et al (2011) Graph theoretical analysis of functional brain networks: test–retest evaluation on short- and long-term resting-state functional MRI data. PLoS One 6:e21976. doi:10.1371/journal.pone.0021976

Weissenbacher A, Kasess C, Gerstl F et al (2009) Correlations and anticorrelations in resting-state functional connectivity MRI: a quantitative comparison of preprocessing strategies. Neuroimage 47:1408–1416. doi:10.1016/j.neuroimage.2009.05.005

Windischberger C, Langenberger H, Sycha T et al (2002) On the origin of respiratory artifacts in BOLD-EPI of the human brain. Magn Reson Imaging 20:575–582. doi:10.1016/S0730-725X(02)00563-5

Wong C-K, Zotev V, Misaki M et al (2016) Automatic EEG-assisted retrospective motion correction for fMRI (aE-REMCOR). Neuroimage 129:133–147. doi:10.1016/j.neuroimage.2016.01.042

Yan CG, Cheung B, Kelly C et al (2013a) A comprehensive assessment of regional variation in the impact of head micro-movements on functional connectomics. Neuroimage 76:183–201. doi:10.1016/j.neuroimage.2013.03.004

Yan CG, Craddock RC, Zuo XN et al (2013b) Standardizing the intrinsic brain: towards robust measurement of inter-individual variation in 1000 functional connectomes. Neuroimage 80:246–262. doi:10.1016/j.neuroimage.2013.04.081

Zang Y, Jiang T, Lu Y et al (2004) Regional homogeneity approach to fMRI data analysis. Neuroimage 22:394–400. doi:10.1016/j.neuroimage.2003.12.030

Zhang D, Raichle ME (2010) Disease and the brains dark energy. Nat Rev Neurol 6:15–28

Zhong S, He Y, Gong G (2015) Convergence and divergence across construction methods for human brain white matter networks: an assessment based on individual differences. Hum Brain Mapp 36:1995–2013. doi:10.1002/hbm.22751

Zou QH, Zhu CZ, Yang Y et al (2008) An improved approach to detection of amplitude of low-frequency fluctuation (ALFF) for resting-state fMRI: fractional ALFF. J Neurosci Methods 172:137–141. doi:10.1016/j.jneumeth.2008.04.012

Zuo XN, Kelly C, Adelstein JS et al (2010) Reliable intrinsic connectivity networks: test–retest evaluation using ICA and dual regression approach. Neuroimage 49:2163–2177. doi:10.1016/j.neuroimage.2009.10.080

Supplementary material:
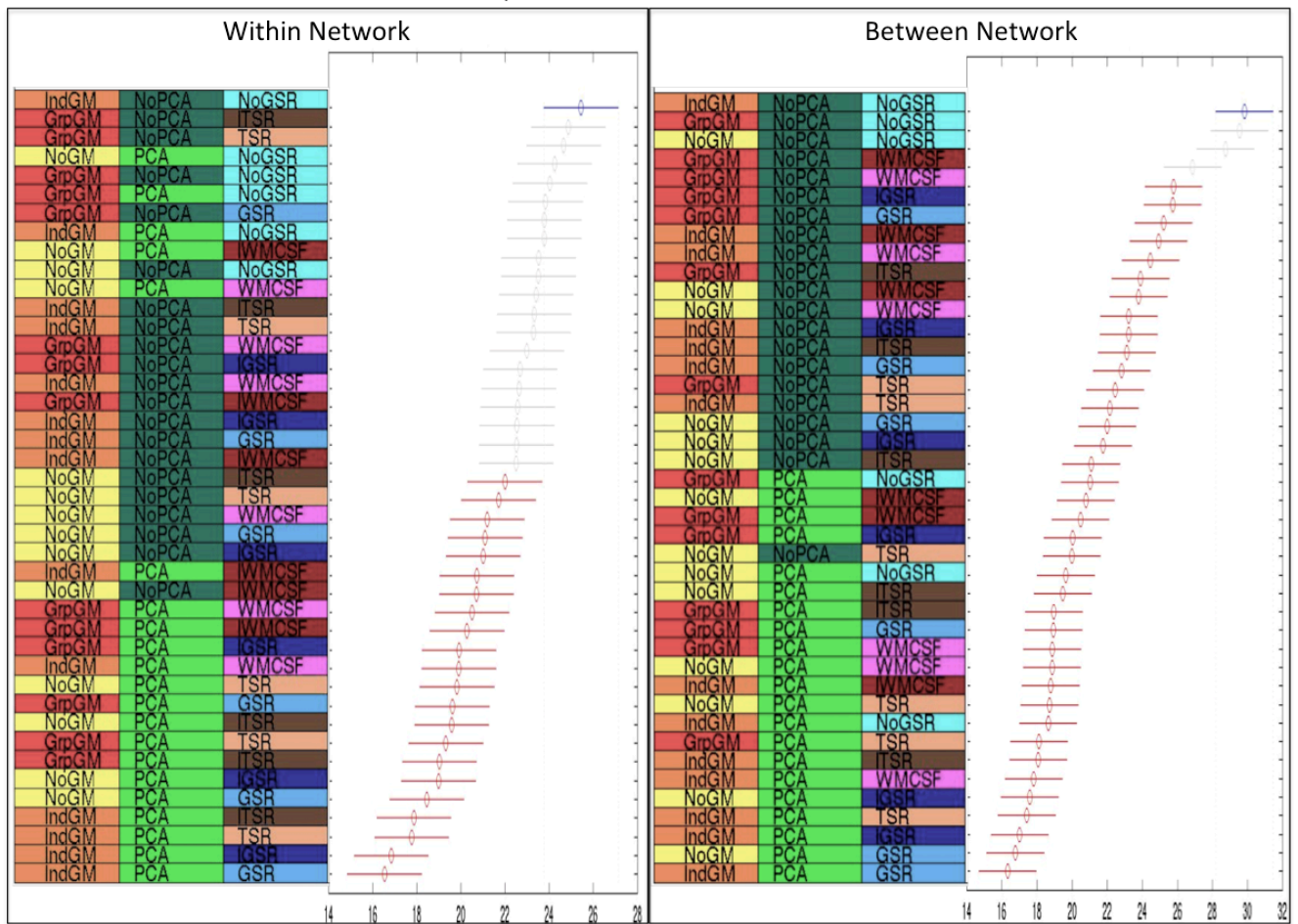
Supplementary figure S1: Individual rankings using a non-parametric Friedman ANOVA of the test-retest reliability on connection level are shown for Spearman correlations of the main analysis conducted using meta-analytically derived cluster volumes.
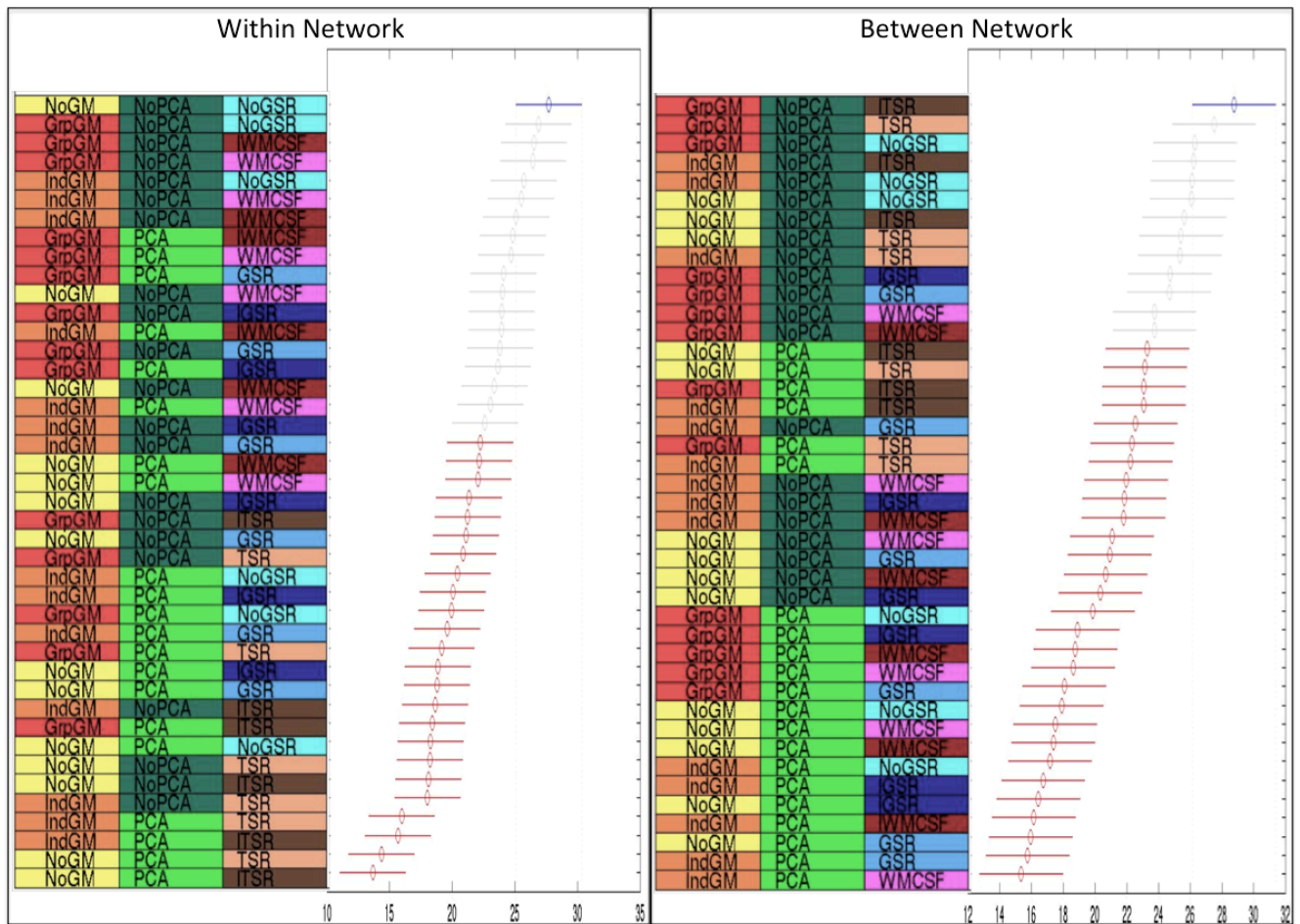
Supplementary figure S2: Individual rankings using a non-parametric Friedman ANOVA of the test-retest reliability on subject level are shown for Spearman correlations of the main analysis conducted using meta-analytically derived cluster volumes.
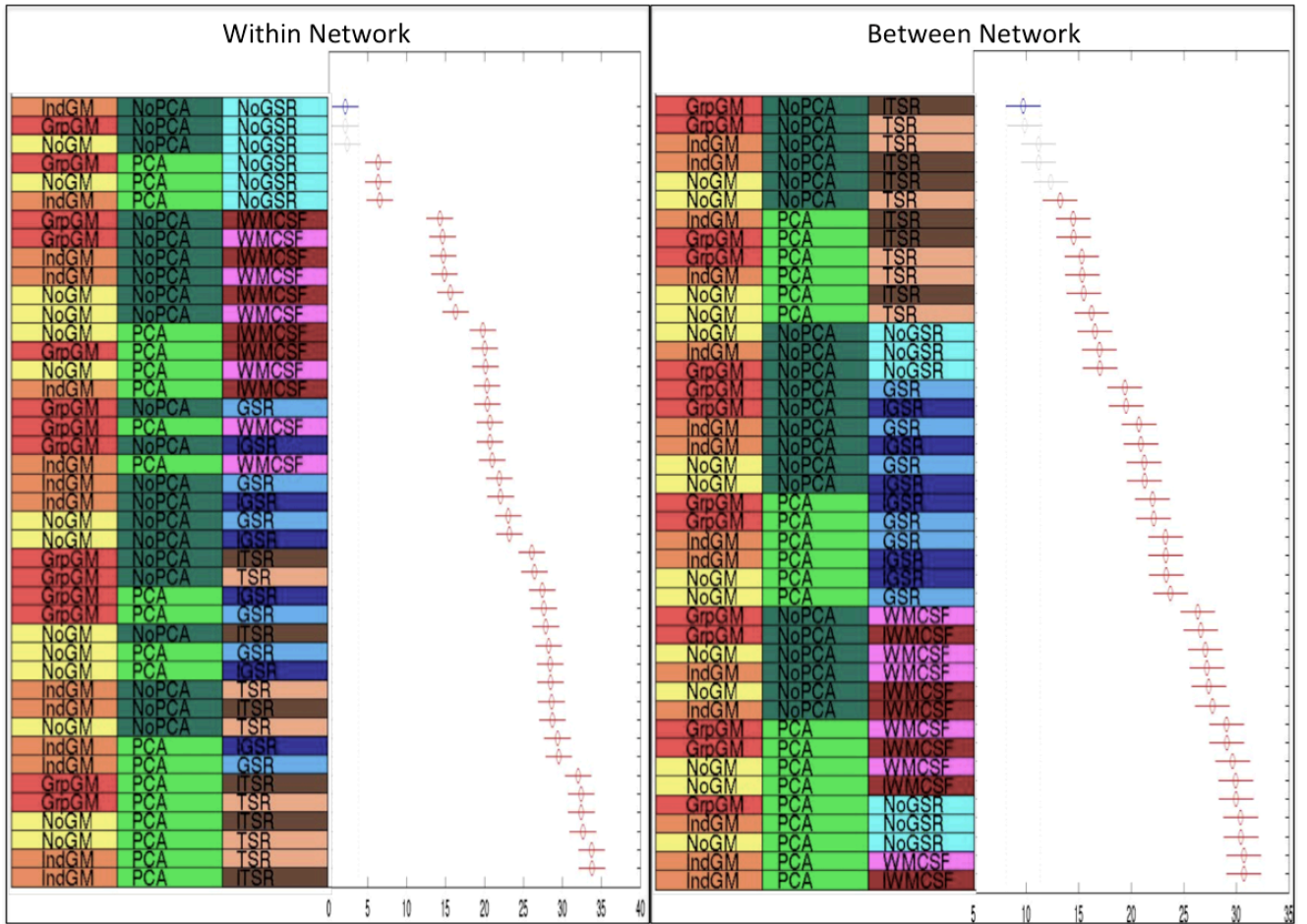
Supplementary figure S3: Individual rankings using a non-parametric Friedman ANOVA of the test-retest reliability on connection level are shown for absolute differences of the main analysis conducted using meta-analytically derived cluster volumes.
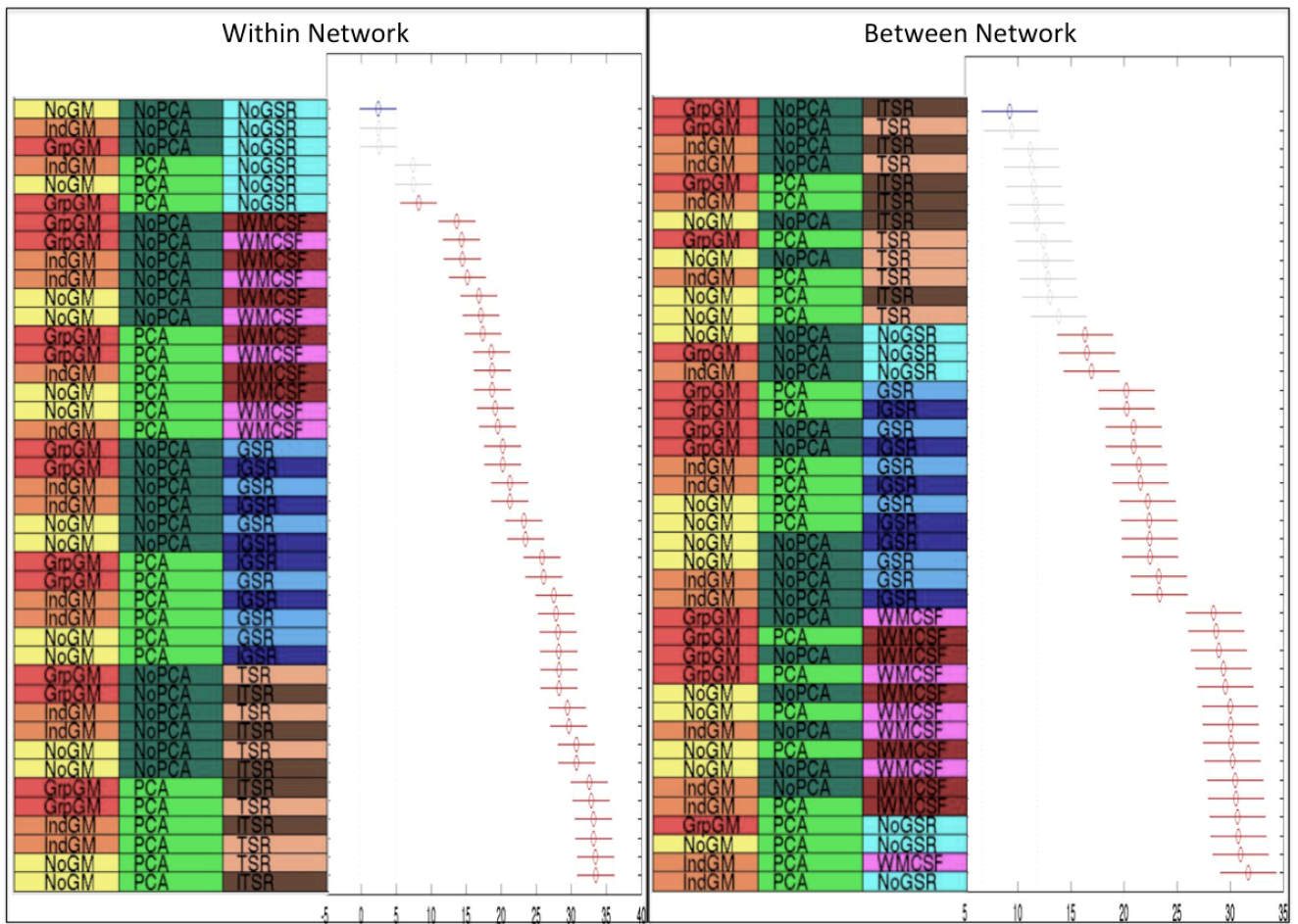
Supplementary figure S4: Individual rankings using a non-parametric Friedman ANOVA of the test-retest reliability on subject level are shown for absolute differences of the main analysis conducted using meta-analytically derived cluster volumes.

Supplementary figure S5: Combined rankings of the test-retest reliability on subject and connection level are shown for Spearman correlations and absolute differences of the analysis conducted using spherical ROIs of 5 mm radius rather than the actual cluster volumes
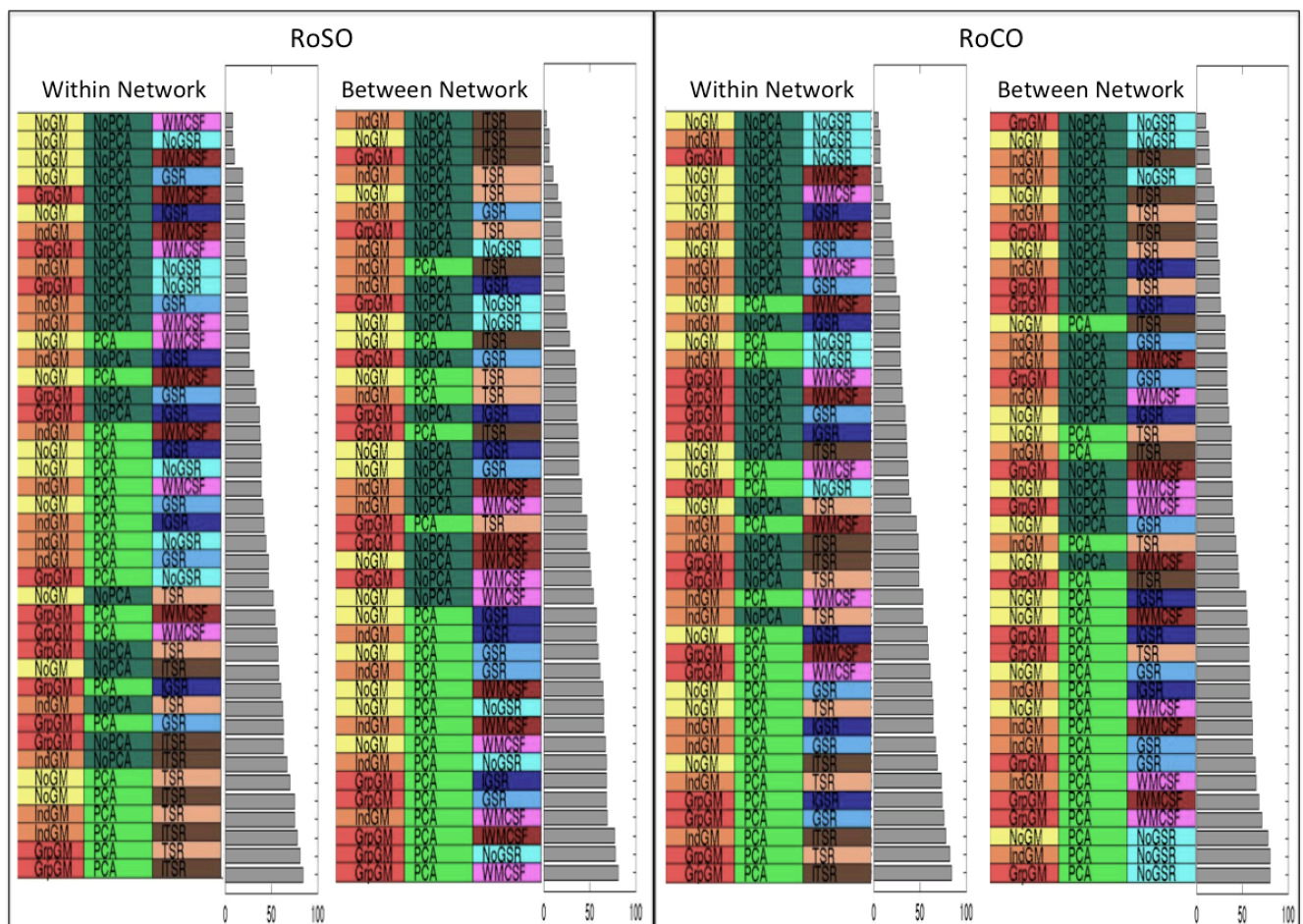
Supplementary figure S6: Combined rankings of the test-retest reliability on subject and connection level are shown for RoSO and RoCO of the analysis conducted using spherical ROIs of 5 mm radius rather than the actual cluster volumes
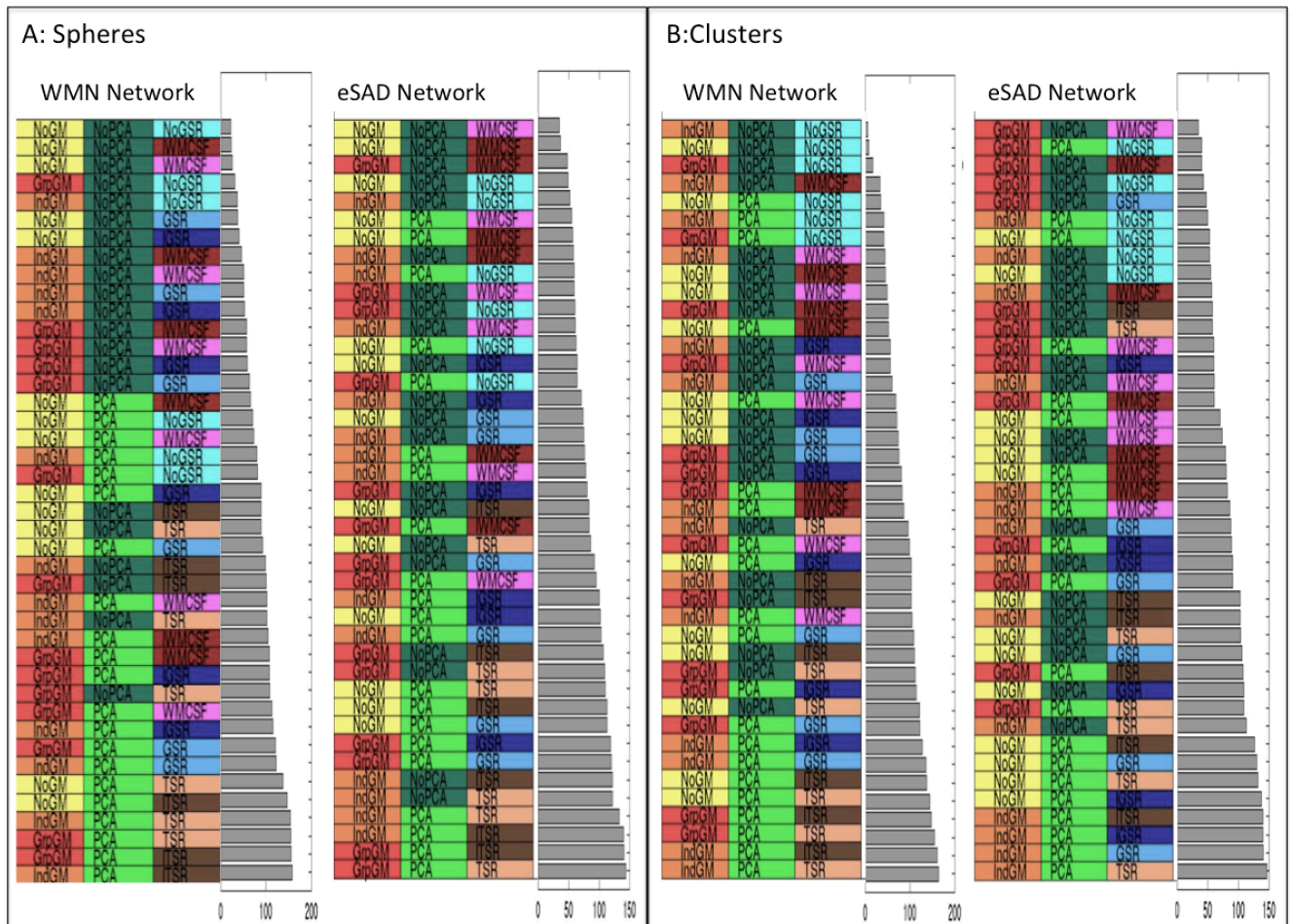
Supplementary figure S7: Summary rankings of reliability across Spearman correlations and absolute differences as well as RoSO and RoCO are shown separately for WMN and eSAD networks for the analyses that defined the volume of interest A) by using spherical ROIs of 5 mm radius and B) by using the actual cluster volumes

Supplementary figure S8: Combined rankings of the test-retest reliability on subject and connection level are shown for Spearman correlations and absolute differences of the analysis conducted using 5 components of PCA derived from the whole brain rather than only the WM and CSF regions.

Supplementary figure S9: Combined rankings of the test-retest reliability on subject and connection level are shown for RoSO and RoCO of the analysis conducted using 5 components of PCA derived from the whole brain rather than only the WM and CSF regions.

Supplementary figure S10: Summary rankings of reliability across Spearman correlations and absolute differences as well as RoSO and RoCO are shown separately for within (WMN & eSAD) and between networks for the analyses A) PCA components derived from the entire brain and B) PCA components derived from the WM & CSF regions.

STUDY 2

On the Integrity of Functional Brain Networks in Schizophrenia, Parkinson's Disease, and Advanced Age: Evidence from Connectivity-Based Single-Subject Classification

Rachel N. Pläschke[1,2,3,4], Edna C. Cieslik[1,2,3,4], Veronika I. Meuller[1,2,3,4], Felix Hoffstaedter[1,2,3,4], Anna Plachti[2,4], Deepthi P. Varikuti[1,2,3,4], Mareike Goosses[4], Anne Latz[1,2,3,4], Svenja Caspers[4,5,6], Christiane Jockwitz[4,5,7], Susanne Moebus[8], Oliver Gruber[9], Claudia R. Eickhoff[2,4,7], Kathrin Reetz[6,10,11], Julia Heller[6,10,11], Martin Seudmeyer[3,12], Christian Mathys[13], Julian Caspers[4,13], Christian Grefkes[14,15], Tobias Kalenscher[16], Robert Langner[1,2,3,4], and Simon B. Eickhoff[1,2,3,4,*]

[1]Institute of Systems Neuroscience, Medical Faculty, Heinrich Heine University Düsseldorf, Düsseldorf, Germany

[2]Institute of Neuroscience and Medicine, Brain & Behavior (INM-7), Research Centre Jülich, Jülich, Germany

[3]Institute of Clinical Neuroscience and Medical Psychology, Medical Faculty, Heinrich Heine University Düsseldorf, Düsseldorf, Germany

[4]Institute of Neuroscience and Medicine, (INM-1), Research Centre Jülich, Jülich, Germany

[5]C. & O. Vogt Institute for Brain Research, Heinrich Heine University Düsseldorf, Düsseldorf, Germany

[6]JARA-BRAIN, Jülich-Aachen Research Alliance, Jülich, Germany

[7]Department of Psychiatry, Psychotherapy and Psychosomatics, RWTH Aachen University, Aachen, Germany

[8]Center for Urban Epidemiology, University of Duisburg-Essen, Essen, Germany

[9]Section for Experimental Psychopathology and Neuroimaging, Department of General Psychiatry, Heidelberg University, Heidelberg, Germany

[10]JARA-BRAIN Institute of Molecular Neuroscience and Neuroimaging (INM-11), Research Centre Jülich, Jülich, Germany

[11]Department of Neurology, RWTH Aachen University, Aachen, Germany

[12]Center for Movement Disorders and Neuromodulation, Department of Neurology, Medical Faculty, Heinrich Heine University Düsseldorf, Düsseldorf, Germany

Own contributions   Data processing
                           Critically revising the paper
                           Total contribution: 10%

On the Integrity of Functional Brain Networks in Schizophrenia, Parkinson's Disease, and Advanced Age: Evidence from Connectivity-Based Single-Subject Classification

Rachel N. Pläschke [iD],[1,2,3,4] Edna C. Cieslik,[1,2,3,4] Veronika I. Müller,[1,2,3,4] Felix Hoffstaedter,[1,2,3,4] Anna Plachti,[2,4] Deepthi P. Varikuti,[1,2,3,4] Mareike Goosses,[4] Anne Latz,[1,2,3,4] Svenja Caspers,[4,5,6] Christiane Jockwitz,[4,5,7] Susanne Moebus,[8] Oliver Gruber,[9] Claudia R. Eickhoff,[2,4,7] Kathrin Reetz,[6,10,11] Julia Heller,[6,10,11] Martin Südmeyer,[3,12] Christian Mathys,[13] Julian Caspers,[4,13] Christian Grefkes,[14,15] Tobias Kalenscher,[16] Robert Langner,[1,2,3,4] and Simon B. Eickhoff[1,2,3,4]*

[1]Institute of Systems Neuroscience, Medical Faculty, Heinrich Heine University Düsseldorf, Düsseldorf, Germany
[2]Institute of Neuroscience and Medicine, Brain & Behavior (INM-7), Research Centre Jülich, Jülich, Germany
[3]Institute of Clinical Neuroscience and Medical Psychology, Medical Faculty, Heinrich Heine University Düsseldorf, Düsseldorf, Germany
[4]Institute of Neuroscience and Medicine, (INM-1), Research Centre Jülich, Jülich, Germany
[5]C. & O. Vogt Institute for Brain Research, Heinrich Heine University Düsseldorf, Düsseldorf, Germany
[6]JARA-BRAIN, Jülich-Aachen Research Alliance, Jülich, Germany
[7]Department of Psychiatry, Psychotherapy and Psychosomatics, RWTH Aachen University, Aachen, Germany
[8]Center for Urban Epidemiology, University of Duisburg-Essen, Essen, Germany
[9]Section for Experimental Psychopathology and Neuroimaging, Department of General Psychiatry, Heidelberg University, Heidelberg, Germany
[10]JARA-BRAIN Institute of Molecular Neuroscience and Neuroimaging (INM-11), Research Centre Jülich, Jülich, Germany
[11]Department of Neurology, RWTH Aachen University, Aachen, Germany
[12]Center for Movement Disorders and Neuromodulation, Department of Neurology, Medical Faculty, Heinrich Heine University Düsseldorf, Düsseldorf, Germany

*Correspondence to: Simon B. Eickhoff, Institute of Neuroscience and Medicine (INM-7), Research Center Jülich, Wilhelm-Johnen-Straße, D-52428 Jülich, Germany. E-mail: S.Eickhoff@fz-juelich.de

[13]Department of Diagnostic and Interventional Radiology, Medical Faculty, Heinrich Heine
University Düsseldorf, Düsseldorf, Germany
[14]Department of Neurology, University Hospital Cologne, Cologne, Germany
[15]Institute of Neuroscience and Medicine, Cognitive Neurology Group (INM-3), Research
Centre Jülich, Jülich, Germany
[16]Comparative Psychology, Heinrich Heine University Düsseldorf, Düsseldorf, Germany

Abstract: Previous whole-brain functional connectivity studies achieved successful classifications of patients and healthy controls but only offered limited specificity as to affected brain systems. Here, we examined whether the connectivity patterns of functional systems affected in schizophrenia (SCZ), Parkinson's disease (PD), or normal aging equally translate into high classification accuracies for these conditions. We compared classification performance between pre-defined networks for each group and, for any given network, between groups. Separate support vector machine classifications of 86 SCZ patients, 80 PD patients, and 95 older adults relative to their matched healthy/young controls, respectively, were performed on functional connectivity in 12 task-based, meta-analytically defined networks using 25 replications of a nested 10-fold cross-validation scheme. Classification performance of the various networks clearly differed between conditions, as those networks that best classified one disease were usually non-informative for the other. For SCZ, but not PD, emotion-processing, empathy, and cognitive action control networks distinguished patients most accurately from controls. For PD, but not SCZ, networks subserving autobiographical or semantic memory, motor execution, and theory-of-mind cognition yielded the best classifications. In contrast, young–old classification was excellent based on all networks and outperformed both clinical classifications. Our pattern-classification approach captured associations between clinical and developmental conditions and functional network integrity with a higher level of specificity than did previous whole-brain analyses. Taken together, our results support resting-state connectivity as a marker of functional dysregulation in specific networks known to be affected by SCZ and PD, while suggesting that aging affects network integrity in a more global way. *Hum Brain Mapp 38:5845–5858, 2017.* © 2017 **Wiley** Periodicals, Inc.

Key words: schizophrenia; Parkinson's disease; normal aging; support vector machine; resting-state fMRI; functional connectivity; brain networks; machine learning

INTRODUCTION

Schizophrenia (SCZ) and Parkinson's disease (PD) are two of the most prevalent and socio-economically relevant brain diseases [Andlin-Sobocki et al., 2005]. Although SCZ onset typically emerges during adolescence and early adulthood [Häfner et al., 2013], PD is characterized by an onset during late adulthood [Hughes et al., 1992; Poewe et al., 2017] and has been associated with premature aging, that is, earlier and more rapid neurodegeneration as compared with the course of normal aging (NA) [Rodriguez et al., 2015]. Both SCZ and PD are characterized by disease-specific pathophysiological changes of the dopaminergic system [Jankovic, 2008; Toda and Abi-Dargham, 2007], contrasting with a more global dopamine decline in NA [Bäckman et al., 2006]. However, it has been proposed that dopaminergic dysfunction in SCZ arises as a secondary effect due to alterations of the glutaminergic system [Laruelle et al., 2003]. In contrast, in PD dopaminergic deficiency represents the primary cause leading to pathophysiological

upstream dysregulations of different neural systems [Obeso et al., 2008]. These neurobiological features of SCZ, PD and NA [Bäckman et al., 2006; Jankovic, 2008; Laruelle et al., 2003; Obeso et al., 2008; Rodriguez et al., 2015; Toda and Abi-Dargham, 2007] may manifest themselves in functional connectivity alterations at the level of large-scale brain networks [Cole et al., 2013; Kelly et al., 2009; Narr and Leaver, 2015; Prodoehl et al., 2014; Sala-Llonch et al., 2015]. However, some putative commonalities (neurodegeneration, dopaminergic dysregulations, and altered connectivity) need to be juxtaposed with the prominent phenotypical differences between SCZ, PD, and NA [Bäckman et al., 2006; Jankovic, 2008; Narr and Leaver, 2015; Prodoehl et al., 2014; Sala-Llonch et al., 2015; Toda and Abi-Dargham, 2007] and the fact that the clinical presentations of SCZ and PD are very different [Eaton et al., 1995; Jankovic, 2008; Kalia and Lang, 2015; van Os and Kapur, 2009], raising the question whether various functional systems are differentially affected in the three conditions. Rather than assessing altered activations in different functional systems by

conducting task-based functional magnetic resonance imaging (fMRI) studies, we examined altered functional connectivity within various functional networks robustly defined by meta-analyses of task-based neuroimaging studies in a comparative fashion [cf. New et al., 2015; Schilbach et al., 2016]. This has the practicable advantage of using easily accessible, short and standardized resting-state (RS) data while at the same time incorporating the consolidated knowledge based on task-based imaging into the analysis. We argue that such an approach is particularly relevant given that in contrast to RS imaging, task-based assessments will rarely be feasible in a routine clinical setting.

Alterations in functional network integrity patterns in SCZ, PD or older adults (compared with respective healthy/young controls) can be captured by using machine learning-based classification. For extracting a diagnostically relevant marker that allows the classification of individual subjects based on the connectivity in functional brain networks, multivariate decoding algorithms like support vector machine (SVM) should provide the most appropriate approach for this endeavor. Rather than testing each connection independently for group differences, SVMs are trained on part of the data by weighting all connections in order to separate the known clinical status from healthy controls (HCs). Classification accuracy can then be determined by assessing the ability to predict group membership of previously unseen subjects. Applied to (whole-brain) connectivity data, this approach has previously been found to distinguish SCZ patients [cf. Arbabshirani et al., 2016; Kambeitz et al., 2015; Wolfers et al., 2015] or PD patients [cf. Chen et al., 2015; Long et al., 2012] from HCs, as well as aged from young subjects (NA) [cf. Meier et al., 2012; Vergun et al., 2013].

Previous pattern-classification studies aimed at providing the best possible classification performance on whole-brain connectivity. In contrast, the aim of this work was to assess whether specific functionally defined networks are altered in SCZ, PD, and NA. Although previous studies mainly used Independent Component Analysis (ICA) based data-driven methods to extract major RS networks [Damoiseaux et al., 2006; Smith et al., 2009], our work is based on a priori meta-analytically defined networks associated with specific sets of behavioral functions such as working memory [Rottschy et al., 2012] or emotional processing [Sabatinelli et al., 2011]. In contrast to well-established RS networks, these networks represent the consolidated information from hundreds of task-based fMRI studies and hence those locations in the brain that are reliably activated when subjects perform tasks pertaining to a particular mental function. We thus argue that these nodes define robust functional networks in the brain related to specific mental domains. In turn, the functions associated with RS networks are usually derived from a reverse inference approach, as these lack any direct relationship to mental functions [Poldrack, 2011]. We suggest that this more direct relationship between the network-nodes and actual task-demands is an important advantage of our approach. Moreover, the employed strategy results in

an a priori, unbiased definition of the respective networks, whereas ICA-based networks are usually defined from the current data [Cole et al., 2010]. Our meta-analytically derived network model approach thus offers the potential to investigate functional connectivity within robust a priori brain networks that are implicated in processing a specific mental process.

Therefore, this study aimed to examine whether the known impairment of different functions in SCZ, PD, or aging, respectively, would equally translate into a high classification accuracy for a given network in the respective group, based on the connectivity pattern within this network. As a "proof-of-principle" approach we therefore intended to investigate whether various a priori networks based on task-activation findings carry differential disease-related information assessable by RS imaging. To this end, we examined two diseases which are clinically very disparate but well studied in the previous neuroimaging literature. The findings were then juxtaposed to findings on age-related effects in the same networks. Thereby, we could evaluate whether the respective networks carry differential information related to the different conditions or, conversely, whether the different networks carry differential information related to a particular condition. Given some putative commonalities and especially phenotypical differences, the aim was to examine the possibility for differential classification of SCZ, PD, und age, rather than to primarily study the specific diseases and their clinical separation from each other or aging per se. In our investigation, these three groups thereby serve as examples to evaluate this approach. For example, we assume that connectivity in the reward (Rew) network will be potent in differentiating SCZ patiens from matched HCs, as several studies have shown impairments related to reward learning in SCZ, and the neurobiology of this network has been linked to psychosis [Deserno et al., 2013; Heinz and Schlagenhauf, 2010; Radua et al., 2015]. Likewise, we would expect a good classification accuracy for PD patients based on FC in the motor network, given that motor impairments represent the core feature of this disease [Jankovic, 2008], and motor circuits in the brains of PD patients are altered during motor tasks and at rest [Herz et al., 2014; Prodoehl et al., 2014; Tessitore et al., 2014]. Finally, NA is accompanied by cognitive decline in various domains [Glisky, 2007], such as deterioration in working memory function [Braver and West, 2008]. For the latter, age-related neural changes have repeatedly been shown at task [Dennis and Cabeza, 2008; Rajah and D'Esposito, 2005] and rest [Keller et al., 2015]. Accordingly, we assume that the working memory (WM) network allows a clear distinction between old and young adults.

In an explorative manner, we furthermore assessed a broad set of networks associated with different behavioural domains (cognitive, social-affective, motivational, and motor-related) since all three conditions (PD, SCZ, and NA) show alterations in various functional domains on the

behavioral and neural level [Barch, 2005; Duncan et al., 2013; Seidler et al., 2010]. Importantly, in our approach, we reasoned that classification performance may be interpreted as an indication for the amount of information contained in a given network regarding a particular disease or age status, and thus of the degree of change observed in the integrity of particular networks under these conditions.

We assume that classification performance will be best for connectivity in those networks that subserve mental functions known to be affected in SCZ and PD. SCZ is characterized by prominent social-affective/motivational alterations [Brunet-Gouet and Decety, 2006; Deserno et al., 2013; Heinz and Schlagenhauf, 2010; Kring and Elis, 2013; Radua et al., 2015], whereas in PD motor impairments are most affected [Herz et al., 2014; Rowe and Siebner, 2012; Tessitore et al., 2014]. We, therefore, hypothesized that social-affective/motivational and motor-related networks provide a superior classification of SCZ and PD patients, respectively. As both diseases are accompanied by cognitive impairments as well, we assumed that cognitive networks may also be predictive to some degree [Barch, 2005; Duncan et al., 2013; Elgh et al., 2009; Nieoullon, 2002]. As NA is associated with a broad spectrum of decline affecting various functional systems (albeit to a varying degree) [Hedden, 2007; Mather, 2016; Seidler et al., 2010], we expected that most networks allowed for an accurate discrimination of old from young adults.

MATERIALS AND METHODS

Samples

Schizrenia

RS fMRI data and phenotypical information of 86 SCZ patients and 84 HCs obtained from the COBRE sample (http://fcon_1000.projects.nitrc.org/indi/retro/cobre.html) and the University Hospital of Göttingen, Germany, were included in the analysis. SCZ diagnosis was assigned as assessed by the DSM-IV-TR based on the structured clinical interview (SCID-P) and the International Classification of Diseases (ICD-10), respectively. SCZ symptom severity was assed using the Positive and Negative Symptom Scale (PANSS) [Kay et al., 1987] evaluating the severity of positive and negative symptoms as well as the general psychopathology. Patients received their regular medication therapy with considerable variability in the exact compounds used and a high prevalence of combination drug therapy (medicated patients but exact medication and dose unknown for Olanzapine equivalent dose [Gardner et al., 2010]: COBRE: 50.9%; Göttingen: 25.8%; medication status unknown: COBRE: 1 SCZ patient; Göttingen: 2 SCZ patients).

Parkinson's disease

RS fMRI data of 80 PD patients and 84 HCs obtained from the RWTH Aachen University Hospital and the University Hospital Düsseldorf, Germany, were included in the analysis. Diagnosis of PD was assigned by consultant neurologists with longstanding expertise in movement disorders based on clinical examination and review of the medical history. Included PD patients fulfilled the standard UK Brain Bank criteria for PD and had on average a mild cognitive impairment as confirmed by the Montreal Cognitive Assessment (MoCA) but no major depression symptoms [Hoops et al., 2009; Hughes et al., 1992; Nasreddine et al., 2005].

To assess PD symptom severity and evaluate motor impairments the Unified Parkinson's Disease Rating Scale Part III [Movement Disorder Society Task Force on Rating Scales for Parkinson's Disease, 2003] (UPDRS) and Hoehn and Yahr Scale (H & Y Scale) [Hoehn and Yahr, 1967] were applied. All patients were medicated with their regular individual PD-related treatment (medication and dose unknown for Levodopa equivalent daily dose [Tomlinson et al., 2010]: Aachen: 28.1%; Düsseldorf: 12.5%).

Healthy controls

RS fMRI data of HC (HC_{SCZ} and HC_{PD}) were obtained from the four different sites as respective clinical subjects (SCZ and PD), and were without any record of neurological or psychiatric disorders as confirmed via structured clinical screening.

Normal aging

RS fMRI data of 95 old (age range: 55–70 years) and 93 young (age range: 20–35 years) participants with an age range of 15 years in each group were obtained from the population-based 1000BRAINS study [Caspers et al., 2014] and another separate study at the Research Centre Jülich, Germany. This relative small age-range aims to enhance the subsample homogeneity. "NA" in old participants refers to the absence of neurodegenerative diseases. Older adults showed cognitive performance adequate for their age (DemTect > 13) as assessed by the Mild Cognitive Impairment and Early Dementia Detection (DemTect) assessment [Kalbe et al., 2004] and all participants did not exhibit clinically relevant symptoms for depression (BDI-II < 13) as evaluated via the Beck Depression Inventory-II [Beck et al., 1996].

Importantly, target and control groups (i.e., patients vs. HCs, old vs. young adults) of all three samples (PD, SCZ, NA) represent subsamples from larger samples that were post-hoc matched for gender, within-scanner movement and (only for the clinical samples) age (cf. Table I for sample and group matching characteristics). Written informed consent from all subjects and approval by the local ethics committees was obtained from all sites. Joint reanalysis of the anonymized data was approved by the ethics committee of the Heinrich Heine University Düsseldorf.

TABLE I. Sample and group matching characteristics

Sample	n (males)	Age (years)	Head movement (DVARS)	Age at onset (years)	Illness duration (years)	Antipsychotic/ dopaminergic medication	Neuropsychology and psychopathology		
SCZ sample						**OZP-equivalent**	**PANSS: Total/PS/NS/GEN**		
COBRE									
SCZ patients	55 (46)	38 ± 14	1.66 ± 0.55*	20 ± 8	17 ± 14	13 ± 8	58 ± 14/14 ± 5/14 ± 5/29 ± 8		
HC$_{scz}$	55 (42)	38 ± 12	1.44 ± 0.41						
Göttingen									
SCZ patients	31 (25)	32 ± 10	1.47 ± 0.30*	25 ± 8	7 ± 8	14 ± 9	52 ± 11/12 ± 3/13 ± 4/28 ± 6		
HC$_{scz}$	29 (22)	32 ± 9	1.31 ± 0.23						
Total									
SCZ patients	86 (71)	36 ± 13	1.59 ± 0.48*						
HC$_{scz}$	84 (64)	36 ± 11	1.39 ± 0.36						
PD sample						**LEDD**	**H & Y Scale**	**UPDRS-III**	**MoCA**
Aachen									
PD patients	32 (21)	64 ± 9	0.51 ± 0.16	59 ± 8	6 ± 5	449 ± 238	2 ± 1	23 ± 12	27 ± 2
HC$_{PD}$	33 (20)	63 ± 6	0.62 ± 0.29						
Düsseldorf									
PD patients	48 (30)	59 ± 9	0.69 ± 0.26	51 ± 9	8 ± 6	1029 ± 416	2.5 ± 1	16 ± 8	24 ± 4
HC$_{PD}$	51 (30)	57 ± 9	0.68 ± 0.22						
Total									
PD patients	80 (51)	61 ± 9	0.62 ± 0.24						
HC$_{PD}$	84 (50)	59 ± 8	0.66 ± 0.25						
NA sample							**DemTect**		**BDI-II**
Jülich									
Old	48 (26)	61 ± 5	1.58 ± 0.41*				16 ± 2		5 ± 5
Young	52 (26)	26 ± 3	1.24 ± 0.24						5 ± 4
1000BRAINS Jülich									
Old	47 (25)	64 ± 4	1.79 ± 0.43*				15 ± 2		6 ± 5
Young	41 (23)	28 ± 4	1.28 ± 0.26						4 ± 4
Total									
Old	95 (51)	63 ± 5	1.68 ± 0.43*						
Young	93 (49)	27 ± 4	1.26 ± 0.25						

SCZ, schizophrenia; HC$_{SCZ}$, matched healthy controls (HCs) of SCZ sample; PD, Parkinson's disease; HC$_{PD}$, matched HCs of PD sample; NA, normal aging; characteristic values in mean ± standard deviation; DVARS, derivative of root mean squared variance over voxels (head movement parameter) [Power et al., 2012]; significant difference in age (clinical samples), gender and movement are marked with * for $P < 0.05$; SCZ: OZP-equivalent [Gardner et al., 2010], Olanzapine equivalent dose; PANSS, Positive and Negative Symptom Scale, (PS, Positive Symptoms Scale/NS, Negative Symptoms Scale/GEN, General Psychopathology Scale); PD: LEDD [Tomlinson et al., 2010], Levodopa equivalent daily dose; H & Y Scale, Hoehn and Yahr Scale; UPDRS-III, Unified Parkinson's Disease Rating Scale Part III; MoCA, Montreal Cognitive Assessment; NA: DemTect, Mild Cognitive Impairment and Early Dementia Detection, BDI-II, Beck Depression Inventory II.

RS fMRI Data Acquisition, Preprocessing, and Analysis

During image acquisition (see Supporting Information Table SI for fMRI parameters), participants were instructed to lie still, let their mind wander and not fall asleep (confirmed at debriefing). SPM8 (www.fil.ion.ucl.ac.uk/spm) was used for image realignment, spatial normalization to the MNI-152 template using the unified segmentation approach [Ashburner and Friston, 2005], and smoothing "5-mm full-width at half-maximum Gaussian kernel".

We investigated 12 functional networks, robustly defined by previous quantitative meta-analyses, to reflect neural correlates of a broad set of cognitive, social-affective/ motivational and motor functions (see Table II for an overview and Supporting Information Table SII for detailed network coordinates and corresponding brain regions). Only meta-analytic networks with a minimum of 10 nodes were included, since a lower number of features are uninformative for robust classification. RS functional connectivity (RSFC) within each network was computed per subject by first extracting the time-series for each node within 6 mm of the meta-analytic peaks. To reduce spurious correlations, variance explained by the six movement parameters and their derivatives (modeled as first and second order effects) as well as the mean white-matter and cerebrospinal fluid signal time-courses was removed from the time series [Satterthwaite et al., 2013; Varikuti et al., 2016]. Subsequently,

TABLE II. Network overview

Network (Abbr.)	Network	Contrast	Nodes	Publications
EmoSF	emotional scene and face processing	emotional scene > neutral scene and emotional face > neutral face	24	Sabatinelli et al. [2011]
ER	cognitive emotion regulation	reappraise > naturalistic emotional responses	14	Buhle et al. [2014]
ToM	theory-of-mind cognition	ToM > non-social baseline	15	Bzdok et al. [2012]
Empathy	empathic processing	"feel into" affect-laden social situations > watched or listened passively	19	Bzdok et al. [2012]
Rew	reward-related decision making	ME: reward valence and decision stages	25	Liu et al. [2011]
AM	autobiographical memory	autobiographical memory > non-autobiographical baseline	22	Spreng et al. [2009]
SM	semantic memory	access to word meaning > processing word structure	23	Binder et al. [2009]
WM	working memory	ME: n-back, sternberg, delayed matching to sample and delayed simple matching tasks	23	Rottschy et al. [2012]
CogAC	cognitive action control	ME: stroop-task, spatial interference task, stop-signal task and go/no-go task	19	Cieslik et al. [2015]
VigAtt	vigilant attention	ME: detection task, discrimination task	16	Langner and Eickhoff [2013]
MNS	mirror neuron system	action observation ∩ action imitation	11	Caspers et al. [2010]
Motor	motor execution	finger tapping > baseline; excl. regions associated with visually paced finger-tapping tasks	10	Witt et al. [2008]

ME, main effect.

time series were high-pass filtered retaining frequencies above 0.01 Hz. Connectivity was computed as the Fisher's Z-transformed Pearson correlation between the time series of each network's nodes; connectivity values were adjusted for effects of acquisition site, gender, movement, total brain volume, and (only for the clinical samples) age [cf. Schilbach et al., 2014, 2016] to avoid classification based on spurious between-subject effects.

SVM Features and Classification

To examine whether the RSFC pattern of a network contains predictive information on the respective groups (SCZ vs. HC$_{SCZ}$, PD vs. HC$_{PD}$, old vs. young) non-sparse linear two-class SVMs were computed using LibSVM [Chang and Lin, 2011] (https://www.csie.ntu.edu.tw/~cjlin/libsvm). SVMs' were trained separately for each of all three analyses (PD, SCZ, NA) and each of the functional networks. Of note, we did not attempt between-patient classification (i.e., PD vs. SCZ), as the different groups were closely matched to their respective controls but substantially different from each other with respect to age, gender, and movement. The input variables (features) to the SVM consisted of edge-wise RSFC between all nodes of a given network. Each SVM was trained and tested by a nested 10-fold cross-validation scheme for each individual group (see e.g., Fig. 1 [Xia et al., 2013]) [cf. Lemm et al., 2011]. The inner loop used a 10-fold cross-validation within the training group to optimize the soft-margin slack parameter. For each fold of the outer loop, the left-out (unseen) 10% were then classified using the SVM trained on the (entire) training-set using the optimized parameter. This nested scheme ensured that classifier optimization and evaluation was performed independent of each other [Kriegeskorte et al., 2009]. Classification performance was evaluated based on accuracy (Acc.) balanced accuracy (bAcc.), sensitivity (Sens.), and specificity (Spec.) as well as two measures derived from signal-detection theory: the area under the receiver operating characteristics (ROC) curve (AUC) [Fawcett, 2004] and d'. Acc. denotes the overall proportion of subjects correctly classified as patients (PD, SCZ) or advanced age versus healthy or younger age, respectively. The bAcc. is calculated as the average proportion of subjects correctly classified as patients (PD, SCZ) or advanced age versus healthy or younger age, respectively. Sens. indicates the percentage of patients (SCZ or PD) correctly classified as ill or subjects correctly classified as old in the aging sample (true positives). Spec. in turn represents the fraction of HCs correctly classified as healthy or subjects correctly identified as young in the aging sample (true negatives). AUC refers to the area under the ROC curve. An ROC curve depicts the relationship between true positive rate and false positive rate, and its AUC value indicates the sensitivity of the diagnostic process independent of any specific decision criterion. Finally, we assessed d', an alternative index of diagnostic sensitivity independent of the decision criterion, calculated
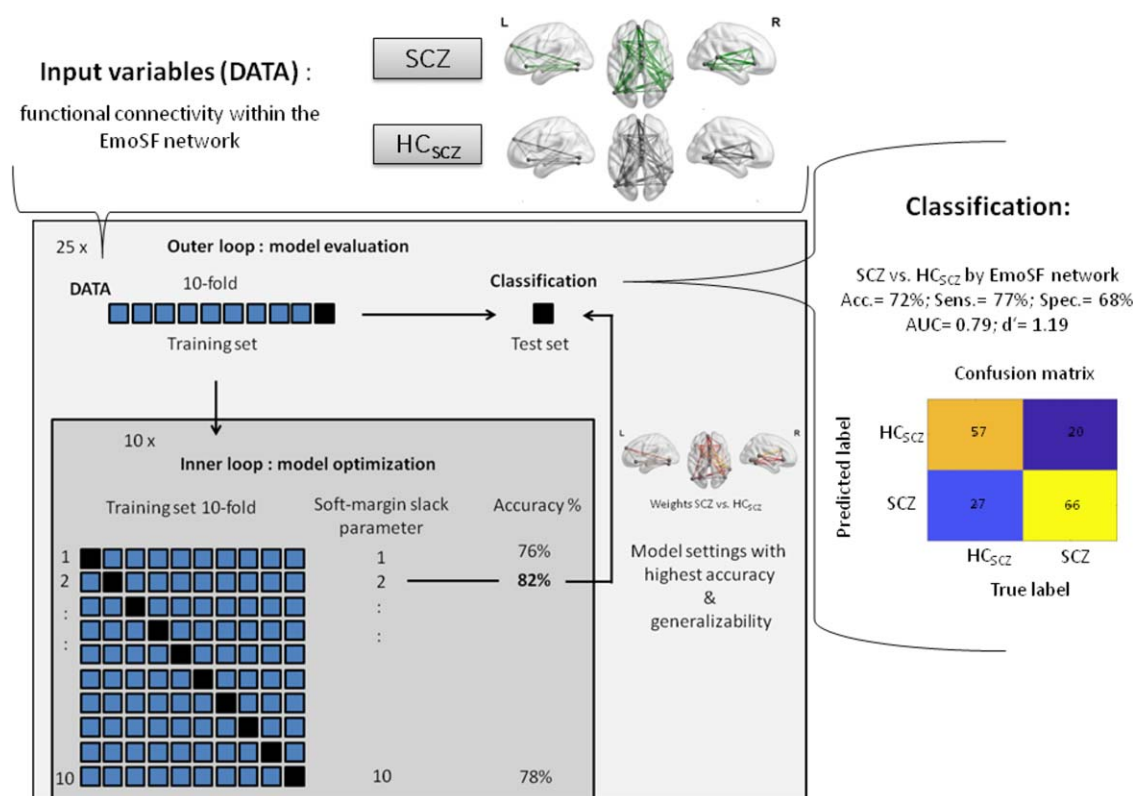
Figure 1.

Linear two-class SVM nested 10-fold cross-validation scheme. Illustration of a SVM example for classification of the SCZ sample based on the EmoSF network. As input variables (DATA) (= features) served the subjects' RSFCs of all edges of every network. The inner loop was performed in a 10-fold manner with 10 repetitions conducted as parameter setting optimization on a training sample. The outer loop was performed in a 10-fold manner with 25 repetitions conducted as classification accuracy testing on an unseen test set. Classification performance measures are computed based on the confusion matrix. Acc., accuracy; Sens., sensitivity; Spec., specificity; AUC, area under the ROC curve and d' (see "Materials and Methods" section for explanation). [Color figure can be viewed at wileyonlinelibrary.com]

as z(true positive rate) – z(false positive rate). To increase robustness, the entire procedure was repeated 25 times, and each performance measures was averaged across repetitions. To examine significant differences in classification performance between networks within each group, pairwise t-tests were performed for each of the 12 networks based on the accuracies obtained from the 25 cross-validation outer loop replications of the separate SVMs (significance threshold of $P < 0.05$, Bonferroni-corrected for the number of pairwise network comparisons).

To compare the separately conducted classifications for SCZ versus HC_{SCZ} and PD versus HC_{PD} subgroups, accuracies obtained for each individual analysis for every network were converted to standardized z-scores by reference to the binomial distribution reflecting chance level and corrected for multiple comparisons by the amount of networks-based classifications. Log-likelihood ratios were estimated to identify networks showing better classification performance for one patient group than the other. To investigate significant differences in classification performance between the groups, t-tests were calculated based on the 25 accuracies obtained from the cross-validation outer loop replications of the separate SVMs performed in each group (SCZ, PD, NA) for each of the 12 networks (significance threshold of $P < 0.05$, Bonferroni-corrected for the number of groups and networks).

RESULTS

As expected, SCZ patients could be distinguished above chance from matched HCs based on RSFC in the Rew network (Acc. = 68%; AUC = 0.73). In turn, PD patients were distinguished above chance from their matched HCs based on RSFC in the motor network (Motor; Acc. = 70%; AUC = 0.77). Finally, old and young subjects were differentiated very well from each other based on RSFC in the WM network (Acc. = 79%; AUC = 0.84). Results are
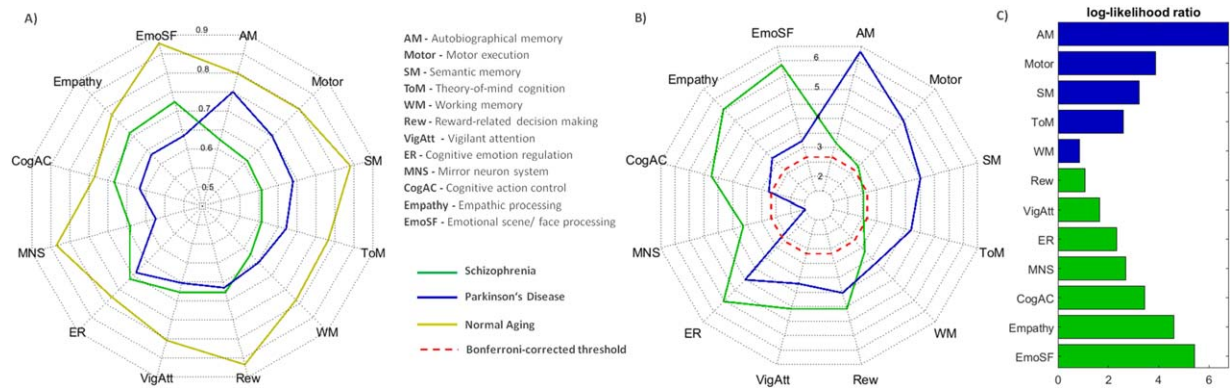
Figure 2.

Group classification results of the SVM. **(A)** Polar plot of group classification accuracies based on all 12 networks for SCZ (in green), PD (in blue) and NA (in yellow). Accuracy refers to the proportion of subjects correctly classified as patients (PD, SCZ) or older age and subjects correctly classified as being HCs or younger age. **(B)** Polar plot of z-standardized accuracies (corrected for multiple comparisons) of patients classification for SCZ (in green) and PD (in blue). **(C)** Log-likelihood ratios of classification performance for networks showing higher classification for one patient group vs. the other. [Color figure can be viewed at wileyonlinelibrary.com]

summarized as follows: Figure 2A for polar plot of group classification accuracies, Table III for Acc., Sens., Spec. and AUC, Supporting Information Table SIII for bAcc., Supporting Information Table SIV for d′, Supporting Information Figure S1 for z-standardized accuracies of all groups and Supporting Information Figure S2 for variance of accuracies.

Considering the performance of all functional networks in distinguishing SCZ and PD patients from their respective HCs, a clear differentiation between networks becomes evident, even though only 2 (SCZ) and 1 (PD) out of 12 networks, respectively, did not significantly exceed chance accuracy (Fig. 2B). The following results and discussion are focused on networks with superior classification performance for the respective disorders. In this context, we would like to re-iterate that we did not attempt to train any classifier to distinguish SCZ from PD patients, since the two samples differed substantially from each other in various confounding factors such as age, gender distribution, and within-scanner movement.

TABLE III. Classification results of the SVM of all groups based on specific networks

Network (Abbr.)	SCZ vs. HC_SCZ Acc. (Sens./Spec.) AUC	PD vs. HC_PD Acc. (Sens./Spec.) AUC	Old vs. Young Acc. (Sens./Spec.) AUC
EmoSF	72% (77%/68%) 0.79	63% (64%/63%) 0.68	88% (89%/86%) 0.93
ER	71% (77%/65%) 0.76	69% (74%/64%) 0.74	78% (79%/76%) 0.86
ToM	61% (74%/46%) 0.62[a]	67% (70%/64%) 0.71	78% (77%/80%) 0.84
Empathy	71% (73%/69%) 0.78	63% (61%/65%) 0.69	78% (80%/75%) 0.83
Rew	68% (73%/62%) 0.73	66% (70%/63%) 0.71	87% (85%/88%) 0.93
AM	62% (67%/57%) 0.71	75% (78%/73%) 0.76	80% (80%/80%) 0.89
SM	61% (67%/54%) 0.68[a]	69% (65%/73%) 0.75	84% (85%/83%) 0.90
WM	62% (65%/60%) 0.66	65% (68%/63%) 0.71	79% (80%/77%) 0.84
CogAC	68% (73%/63%) 0.69	62% (66%/57%) 0.67	73% (73%/74%) 0.83
VigAtt	68% (72%/63%) 0.72	65% (68%/63%) 0.67	80% (78%/83%) 0.89
MNS	64% (65%/63%) 0.73	57% (64%/51%) 0.53[a]	84% (83%/84%) 0.91
Motor	61% (72%/50%) 0.61	70% (68%/73%) 0.77	80% (79%/81%) 0.90

Abbreviations: Acc., Accuracy (in %)/Sens., sensitivity (in %)/Spec., specificity (in %)/AUC, area under the ROC curve.
[a]Network with no significant classification result.
Acc. refers to the proportion of subjects correctly classified as patients (PD, SCZ) or older age and subjects correctly classified as being healthy or younger age (mean of sensitivity and specificity). Sensitivity relates to the percentage of patients (SCZ or PD) correctly classified as being ill or else subjects correctly identified as old in the aging sample (true positives). Specificity relates to the percentage of healthy subjects correctly classified as being healthy or else subjects correctly identified as young in the aging sample (true negatives). AUC refers to the area under the ROCs curve. The ROC curve depicts the relationship between true positive rate and false positive rate.

For SCZ, the emotional scene and face processing (EmoSF) network (Acc. = 72%; AUC = 0.79) as well as the networks for empathic processing (Empathy; Acc. = 71%; AUC = 0.78) and cognitive action control (CogAC; Acc. = 68%; AUC = 0.69) distinguished patients most accurately from their HCs. Hence these networks' connectivity patterns may be considered to contain the highest level of information with respect to SCZ. The EmoSF network was significantly better in the SCZ classification compared with all other networks ($P < 0.001$). For PD, the networks subserving autobiographical memory (AM; Acc. = 75%; AUC = 0.76), motor execution (Motor; Acc. = 70%; AUC = 0.77), semantic memory (SM; Acc. = 69%; AUC = 0.75), and theory-of-mind cognition (ToM; Acc. = 67%; AUC = 0.71) yielded the highest classification accuracies, that is, contained the most informative PD-related differences in RSFC. The AM network was significantly better in the PD classification compared with all other networks ($P < 0.001$). All network comparison results within the patient groups are summarized in Supporting Information Tables SV and SVI.

The between-network comparison of classification performance with respect to SCZ and PD revealed that the networks discriminating either disorder from their respective controls were highly specific (Fig. 2B,C), indicating that these networks carry differential amounts of information regarding SCZ and PD, respectively. In particular, both EmoSF and Empathy networks showed the best performance at distinguishing SCZ patients from HCs (EmoSF: $z = 5.9$; Empathy: $z = 5.5$) but were notably worse at discriminating PD patients from their HCs (EmoSF: $z = 3.2$; Empathy: $z = 3.2$). Similarly, the CogAC network exhibited high accuracy at classifying SCZ patients and their respective HCs ($z = 4.7$) but inferior performance at distinguishing PD patients from their HCs ($z = 2.7$).

In turn, the motor network very well classified PD patients and their HCs ($z = 5$) but was remarkably ineffective at classifying SCZ patients and their HCs ($z = 2.9$). Likewise, the AM and SM networks achieved high accuracies in classifying PD patients and controls (AM: $z = 6.3$; SM: $z = 4.5$) but performed much less well when classifying SCZ patients and controls (AM: $z = 3.2$; SM: $z = 2.5$). Networks which were most accurate in distinguishing SCZ from HCs (EmoSF, Empathy, and CogAC) exhibited significant better classification performance in the SCZ group compared to the PD group (EmoSF: $P < 0.001$; Empathy: $P < 0.001$; CogAC: $P < 0.001$; Supporting Information Table SVII). Likewise, networks which performed best at discriminating PD patients from HCs (AM, Motor, SM, and ToM) showed significant better classification performance in the PD group compared with the SCZ group (AM: $P < 0.001$; Motor: $P < 0.001$; SM: $P < 0.001$; ToM: $P < 0.001$; Supporting Information Table SVII).

This differential picture markedly contrasted with the results obtained for the classification between old and young subjects. In the aging sample, each network yielded accuracies $\geq 73\%$ (see Supporting Information Table SVIII for network comparison results within NA), significantly outperforming every classification obtained in the SCZ or PD samples ($P < 0.001$; see Fig. 2A, Supporting Information Figure S1, Table III, Supporting Information Tables SIX and SX).

In particular, for each network the accuracy for classifying a previously unseen participant as young or old was about 10% higher than any clinical classification based on the same network. Additionally, the comparison of all three separate group classifications revealed that the variance of the classification accuracies over the 25 replications of the outer loop was distinctly lower for the classification of age, as compared with classifying the clinical status (Supporting Information Fig. S2).

DISCUSSION

We assessed whether RSFC patterns in a diverse set of functionally defined brain networks allowed for a classification of patients with SCZ or PD or healthy older adults on the one hand, and their respective healthy or young controls on the other. Thereby, we evaluated which functional system was most informative for a given condition (i.e., SCZ, PD, or higher age). Conversely, our analysis also assessed the amount of information on each condition found in a given network. Our results show in a proof-of-principle manner that networks pertaining to functions known to be affected by SCZ, PD, or aging indeed exhibited good classification performance for the respective condition. Furthermore, each network's young–old classification outperformed any disease-related classification. This indicates that specific networks are affected by and associated with the diseases, whereas for healthy older adults RSFC appears to be altered rather globally.

Conceptual Considerations

Our study demonstrates that machine-learning techniques can be successfully used to assess whether RSFC in functional systems known to be affected in SCZ, PD, or advanced age exhibits high classification capacity for the respective condition. Further, our approach compared the classification capacity of RSFC patterns between different functional networks and between several clinical and physiological states. Of note, for each classification, target and control groups (i.e., SCZ vs. HC_{scz}, PD vs. HC_{PD}, old vs. young) were well matched with respect to gender and (for the clinical samples) age. In addition, RSFC variance attributable to these confounding factors or within-scanner movement was regressed out of the data before the SVM analyses. Therefore, these confounds were evidently heterogeneous across the three groups (SCZ, PD, NA) but should not have influenced classification accuracy within each condition. In spite of proper matching and state-of-the-art removal of variance related to motion [cf. Power

et al., 2012; Satterthwaite et al., 2013], residual effects that only manifest in the multivariate pattern cannot be fully ruled out. However, one factor worth noting is that, for example, we observed differential classification performance across networks in the SCZ sample, largely ruling out a dominant general effect of head motion.

Given that both groups were assessed under their regular medication, differences in classification performance may be influenced by pharmacological treatment. In particular, we cannot exclude that classification results of networks modulated via dopaminergic transmission (e.g., reward or motor system) might originate from interactions between disease condition and medication. Unfortunately, however, we could not perform a more detailed assessment of the influence of medication, as the compounds, duration of treatment and doses varied considerably between subjects, with many receiving a combination of drugs.

When comparing classification performance to previous work based on whole-brain functional connectomes [cf. Chen et al., 2015; Long et al., 2012; Meier et al., 2012; Su et al., 2013; Tang et al., 2012; Vergun et al., 2013; Yu et al., 2013], we note that our approach yielded higher functional specificity, allowing inference on the amount of disease-specific information in well-defined functional systems. We acknowledge that even though most of the classifications well exceeded chance level, the achieved network-based classification accuracies are not strong enough for successful connectivity-based single-subject diagnosis. Still, our "sparse" approach achieved classification accuracies comparable to those reported in previous whole-brain studies, whose feature space obviously was substantially larger than ours. This is particularly noteworthy given that two further aspects besides feature space could be expected to decrease classifier performance in our study [Arbabshirani et al., 2016; Haller et al., 2014; Kambeitz et al., 2015; Schnack and Kahn, 2016; Varoquaux et al., 2016]: First, all of our three groups were based on relatively large samples that were combined from two different measurement sites and hence should be more heterogeneous than usual. Second, we used replicated 10-fold cross-validation, rather than the more optimistic leave-one-out approach [Varoquaux et al., 2016]. We thus argue that the chosen combination of examining robustly defined functional networks and optimized analysis through replicated and nested 10-fold cross-validation may provide valuable new insights into the pathophysiology of brain disorders that is not attainable through global analyses of the entire functional connectome.

Classification of SCZ Patients and Controls

We found that the networks subserving EmoSF, empathic processing as well as CogAC yielded the best performance. Aberrant processing of emotional stimuli [Takahashi et al., 2004] and impaired abilities to relate to others' affective states [Benedetti et al., 2009; Derntl et al.,

2012; Harvey et al., 2012] are features of SCZ and mirrored in the degree of SCZ-related information that is contained in the EmoSF (AUC = 0.79) and Empathy (AUC = 0.78) networks. Further, the good classification performance of the CogAC network resonates well with alterations in cognitive control processes as a core deficit in SCZ [cf. Lesh et al., 2011].

Somewhat surprisingly, the Rew network did not differentiate SCZ from HCs with high accuracy, given the prominent role of the dopaminergic system [Toda and Abi-Dargham, 2007] and aberrant salience processing in psychosis [Heinz and Schlagenhauf, 2010; Radua et al., 2015] and the association with the reward system in this disorder. We conjecture that this lack of predictive information could arise from the fact that in contrast to task-activation data, RSFC analyses primarily capture the tonic rather than phasic state of these networks [Schultz et al., 1997].

Classification of PD Patients and Controls

The superior classification performance observed for the motor execution network (AUC = 0.77) is hardly surprising, since motor impairments represent a key clinical feature of PD, and differences in action-related brain circuitry are well established in this disorder [Herz et al., 2014; Rowe and Siebner, 2012; Tessitore et al., 2014]. The finding that the AM (AUC = 0.76) and SM (AUC = 0.75) networks also achieved a very good differentiation of PD patients from HCs was rather surprising, though. Although PD is a neurodegenerative disorder and dementia is common in PD patients [Aarsland et al., 2001, 2003], several patients showed evidence for mild cognitive impairment, using the MoCA for screening. We can hence only speculate that the RSFC differences in AM and SM networks may pick up these deficits as revealed by standard behavioral screening instruments.

Finally, the good classification performance achieved by the ToM network (AUC = 0.71) was unexpected but matches a growing literature of impaired social cognition in PD patients [Bora et al., 2015; Poletti et al., 2011; Díez-Cirarda et al., 2015].

Age Group Classification

One of the most striking observations from this study was that every single network achieved a better classification with respect to age group than with respect to SCZ or PD. While we hypothesized that the broad spectrum of age-related changes in various mental functions [Craik and Salthouse, 2011; Glisky, 2007; Seidler et al., 2010] would be reflected by changes in several networks [Craik and Salthouse, 2011; Hedden, 2007; Mather, 2016; Seidler et al., 2010], the consistency (across both networks and replications) of high classification accuracies is intriguing. It stands to reason that the mechanisms underlying the discriminative changes in functional connectivity patterns

may be diverse. In particular, they should include neuro-degeneration (cognitive networks [Hedden, 2007]), neuro-chemical changes (Rew networks [Bäckman et al., 2006]), altered affective processing (social-affective networks [Mather, 2016]) and use-dependent plasticity (motor networks [Demirakca et al., 2016]). In addition, it may be argued that in spite of all inter-individual variability age-related changes represent a more homogeneous change of the neuro-functional architecture [Ferreira et al., 2016; Meier et al., 2012] relative to the inevitable heterogeneity among clinical populations.

Given that connectivity patterns of all systems differentiated very well between young and old participants, we acknowledge the possibility that the relevant drivers may be of non-neural origin. In particular, despite of our optimized confound removal [Power et al., 2012; Satterthwaite et al., 2013; Varikuti et al., 2016], we cannot exclude that residual effects related to motion or brain atrophy as well as physiological effects such as macro- and microvascular changes and their cumulative impact on hemodynamic signals [D'Esposito et al., 2003] may have contributed to our findings.

Although the contributions of neural and non-neural effects outlined in this section certainly warrant further investigation, one of the most critical conclusions that should be taken from the high classification accuracy between younger and older participants is the danger of obtaining spuriously high accuracies in clinical classification studies if patients and HCs are not carefully matched for age.

Conclusions and Outlook

We investigated the potential of RS connectivity patterns in a wide variety of functional networks to distinguish SCZ and PD patients from matched HCs as well as old from young adults. We showed that networks defined by robust activation due to mental operations known to be affected in the respective condition indeed contained information on the respective condition that is captured by our pattern-classification approach and translates into good classification accuracies. Classification accuracies obtained through replicated, nested 10-fold cross-validation were not only generally comparable to those obtained from whole-brain analyses but also revealed a differentiated picture for both disorders in comparisons. Both SCZ and PD were specifically well predicted by distinct networks that resonate well with known clinical and pathophysiological features. The presented approach thus opens an avenue toward robust and more specific assessments of clinical and developmental differences in functional systems than previous whole-brain analyses. One of the most striking findings of this work was the fact that integrity in all networks was much better at identifying participants with advanced age than with any of the two disorders. While the most likely heterogeneous mechanisms behind this

phenomenon certainly need to be addressed in more detail, the current findings highlight the importance of considering age-related effects as a potential source of bias in clinical classification studies.

FINANCIAL DISCLOSURES

REFERENCES

Aarsland D, Andersen K, Larsen JP, Lolk A, Nielsen H, Kragh-Sørensen P (2001): Risk of dementia in Parkinson's disease: A community-based, prospective study. Neurology 56:730–736.

Aarsland D, Andersen K, Larsen JP, Lolk A, Kragh-Sørensen P (2003): Prevalence and characteristics of dementia in Parkinson disease: An 8-year prospective study. Arch Neurol 60:387–392.

Andlin-Sobocki P, Jönsson B, Wittchen H-U, Olesen J (2005): Cost of disorders of the brain in Europe. Eur J Neurol 12(Suppl 1): 1–27.

Arbabshirani MR, Plis S, Sui J, Calhoun VD (2016): Single subject prediction of brain disorders in neuroimaging: Promises and pitfalls. NeuroImage 145:137–165.

Ashburner J, Friston KJ (2005): Unified segmentation. NeuroImage 26:839–851.

Bäckman L, Nyberg L, Lindenberger U, Li S-C, Farde L (2006): The correlative triad among aging, dopamine, and cognition: Current status and future prospects. Neurosci Biobehav Rev 30:791–807.

Barch DM (2005): The cognitive neuroscience of schizophrenia. Annu Rev Clin Psychol 1:321–353.

Beck AT, Steer RA, Brown G (1996): Beck Depression Inventory-Second Edition (BDI-II). San Antonio, TX: The Psychological Corporation.

Benedetti F, Bernasconi A, Bosia M, Cavallaro R, Dallaspezia S, Falini A, Poletti S, Radaelli D, Riccaboni R, Scotti G, Smeraldi E (2009): Functional and structural brain correlates of theory of mind and empathy deficits in schizophrenia. Schizophr Res 114:154–160.

Binder JR, Desai RH, Graves WW, Conant LL (2009): Where Is the Semantic System? A Critical Review and Meta-Analysis of 120 Functional Neuroimaging Studies. Cereb Cortex 19:2767–2796.

Bora E, Walterfang M, Velakoulis D (2015): Theory of mind in Parkinson's disease: A meta-analysis. Behav Brain Res 292: 515–520.

Braver TS, West R (2008): Working memory, executive control, and aging. In: Craik, FIM, Salthouse, TA, editors. The handbook of aging and cognition, 3rd ed. New York, NY, US: Psychology Press. pp 311–372.

Brunet-Gouet E, Decety J (2006): Social brain dysfunctions in schizophrenia: A review of neuroimaging studies. Psychiatry Res Neuroimaging 148:75–92.

Buhle JT, Silvers JA, Wager TD, Lopez R, Onyemekwu C, Kober H, Weber J, Ochsner KN (2014): Cognitive reappraisal of emotion: a meta-analysis of human neuroimaging studies. Cereb Cortex 24:2981–2990.

Bzdok D, Schilbach L, Vogeley K, Schneider K, Laird AR, Langner R, Eickhoff SB (2012): Parsing the neural correlates of moral cognition: ALE meta-analysis on morality, theory of mind, and empathy. Brain Struct Funct 217:783–796.

Caspers S, Moebus S, Lux S, Pundt N, Schütz H, Mühleisen TW, Gras V, Eickhoff SB, Romanzetti S, Stöcker T, Stirnberg R, Kirlangic ME, Minnerop M, Pieperhoff P, Mödder U, Das S, Evans AC, Jöckel K-H, Erbel R, Cichon S, Nöthen MM, Sturma D, Bauer A, Jon Shah N, Zilles K, Amunts K (2014): Studying variability in human brain aging in a population-based German cohort—rationale and design of 1000BRAINS. Front Aging Neurosci 6:149.

Caspers S, Zilles K, Laird AR, Eickhoff SB (2010): ALE meta-analysis of action observation and imitation in the human brain. NeuroImage 50:1148–1167.

Chang C, Lin C-J (2011): LIBSVM. A library for support vector machines. ACM Trans Intell Syst Technol 2:1–27:27.

Chen Y, Yang W, Long J, Zhang Y, Feng J, Li Y, Huang B (2015): Discriminative analysis of Parkinson's disease based on whole-brain functional connectivity. PloS One 10:e0124153.

Cieslik EC, Mueller VI, Eickhoff CR, Langner R, Eickhoff SB (2015): Three key regions for supervisory attentional control: evidence from neuroimaging meta-analyses. Neurosci Biobehav Rev 48:22–34.

Cole DM, Beckmann CF, Oei NYL, Both S, van Gerven JMA, Rombouts SARB (2013): Differential and distributed effects of dopamine neuromodulations on resting-state network connectivity. NeuroImage 78:59–67.

Cole DM, Smith SM, Beckmann CF (2010): Advances and pitfalls in the analysis and interpretation of resting-state FMRI data. Front Syst Neurosci 4:8.

Craik FIM, Salthouse TA (2011): The Handbook of Aging and Cognition, 3rd ed. New York, NY: Psychology Press.

Damoiseaux JS, Rombouts SARB, Barkhof F, Scheltens P, Stam CJ, Smith SM, Beckmann CF (2006): Consistent resting-state networks across healthy subjects. Proc Natl Acad Sci U S A 103:13848–13853.

Demirakca T, Cardinale V, Dehn S, Ruf M, Ende G (2016): The exercising brain: Changes in functional connectivity induced by an integrated multimodal cognitive and whole-body coordination training. Neural Plast 2016:8240894.

Dennis NA, Cabeza R (2008): Neuroimaging of healthy cognitive aging. In: Craik, FIM, Salthouse, TA, editors. The Handbook of Aging and Cognition, 3rd ed. New York, NY: Psychology Press. pp 1–54.

Derntl B, Finkelmeyer A, Voss B, Eickhoff SB, Kellermann T, Schneider F, Habel U (2012): Neural correlates of the core facets of empathy in schizophrenia. Schizophr Res 136:70–81.

Deserno L, Boehme R, Heinz A, Schlagenhauf F (2013): Reinforcement learning and dopamine in schizophrenia: Dimensions of symptoms or specific features of a disease group?. Front Psychiatry 4:172.

D'Esposito M, Deouell LY, Gazzaley A (2003): Alterations in the BOLD fMRI signal with ageing and disease: a challenge for neuroimaging. Nat Rev Neurosci 4:863–872.

Díez-Cirarda M, Ojeda N, Peña J, Cabrera-Zubizarreta A, Gómez-Beldarrain MÁ, Gómez-Esteban JC, Ibarretxe-Bilbao N (2015): Neuroanatomical correlates of theory of mind deficit in Parkinson's disease: A Multimodal Imaging Study. PLoS One 10: e0142234.

Duncan GW, Firbank MJ, O'Brien JT, Burn DJ (2013): Magnetic resonance imaging: A biomarker for cognitive impairment in Parkinson's disease?. Mov Disord off J Mov Disord Soc 28:425–438.

Eaton WW, Thara R, Federman B, Melton B, Liang KY (1995): Structure and course of positive and negative symptoms in schizophrenia. Arch Gen Psychiatry 52:127–134.

Elgh E, Domellöf M, Linder J, Edström M, Stenlund H, Forsgren L (2009): Cognitive function in early Parkinson's disease: A population-based study. Eur J Neurol 16:1278–1284.

Fawcett T (2004): ROC Graphs: Notes and Practical Considerations for Researchers. Machine Learning 31:1–38.

Ferreira LK, Regina ACB, Kovacevic N, Martin M. D GM, Santos PP, Carneiro C, de G, Kerr DS, Amaro E, McIntosh AR, Busatto GF (2016): Aging effects on whole-brain functional connectivity in adults free of cognitive and psychiatric disorders. Cereb Cortex 26:3851–3865.

Gardner DM, Murphy AL, O'Donnell H, Centorrino F, Baldessarini RJ (2010): International consensus study of antipsychotic dosing. Am J Psychiatry 167:686–693.

Glisky EL (2007): Changes in Cognitive Function in Human Aging. In: Riddle DR, editor. Brain Aging: Models, Methods, and Mechanisms. Boca Raton, FL: CRC Press/Taylor & Francis.

Häfner H, Maurer K, an der Heiden W (2013): ABC Schizophrenia study: An overview of results since 1996. Soc Psychiatry Psychiatr Epidemiol 48:1021–1031.

Haller S, Lovblad K-O, Giannakopoulos P, Van De Ville D (2014): Multivariate pattern recognition for diagnosis and prognosis in clinical neuroimaging: State of the art, current challenges and future trends. Brain Topogr 27:329–337.

Harvey P-O, Zaki J, Lee J, Ochsner K, Green MF (2012): Neural substrates of empathic accuracy in people with Schizophrenia. Schizophr Bull sbs042 39:617–628.

Hedden T (2007): Imaging cognition in the aging human brain. In: Riddle DR, editor. Brain Aging: Models, Methods, and Mechanisms. Boca Raton, FL: CRC Press/Taylor & Francis.

Heinz A, Schlagenhauf F (2010): Dopaminergic dysfunction in schizophrenia: Salience attribution revisited. Schizophr Bull 36: 472–485.

Herz DM, Eickhoff SB, Løkkegaard A, Siebner HR (2014): Functional neuroimaging of motor control in Parkinson's disease: A meta-analysis. Hum Brain Mapp 35:3227–3237.

Hoehn MM, Yahr MD (1967): Parkinsonism: Onset, progression and mortality. Neurology 17:427–442.

Hoops S, Nazem S, Siderowf AD, Duda JE, Xie SX, Stern MB, Weintraub D (2009): Validity of the MoCA and MMSE in the detection of MCI and dementia in Parkinson disease. Neurology 73:1738–1745.

Hughes AJ, Daniel SE, Kilford L, Lees AJ (1992): Accuracy of clinical diagnosis of idiopathic Parkinson's disease: A clinico-pathological study of 100 cases. J Neurol Neurosurg Psychiatry 55:181–184.

Jankovic J (2008): Parkinson's disease: Clinical features and diagnosis. J Neurol Neurosurg Psychiatry 79:368–376.

Kalbe E, Kessler J, Calabrese P, Smith R, Passmore AP, Brand M, Bullock R (2004): DemTect: A new, sensitive cognitive screening test to support the diagnosis of mild cognitive impairment and early dementia. Int J Geriatr Psychiatry 19:136–143.

Kalia LV, Lang AE (2015): Parkinson's disease. Lancet Lond Engl 386:896–912.

Kambeitz J, Kambeitz-Ilankovic L, Leucht S, Wood S, Davatzikos C, Malchow B, Falkai P, Koutsouleris N (2015): Detecting

neuroimaging biomarkers for schizophrenia: A meta-analysis of multivariate pattern recognition studies. Neuropsychopharmacol off Publ Am Coll Neuropsychopharmacol 40:1742–1751.

Kay SR, Flszbein A, Opfer LA (1987): The positive and negative syndrome scale (PANSS) for schizophrenia. Schizophr. Bull 13: 261–276.

Keller JB, Hedden T, Thompson TW, Anteraper SA, Gabrieli JDE, Whitfield-Gabrieli S (2015): Resting-state anticorrelations between medial and lateral prefrontal cortex: Association with working memory, aging, and individual differences. Cortex J Devoted Study Nerv Syst Behav 64:271–280.

Kelly C, de Zubicaray G, Di Martino A, Copland DA, Reiss PT, Klein DF, Castellanos FX, Milham MP, McMahon K (2009): L-dopa modulates functional connectivity in striatal cognitive and motor networks: A double-blind placebo-controlled study. J Neurosci 29:7364–7378.

Kriegeskorte N, Simmons WK, Bellgowan PSF, Baker CI (2009): Circular analysis in systems neuroscience: The dangers of double dipping. Nat Neurosci 12:535–540.

Kring AM, Elis O (2013): Emotion deficits in people with Schizophrenia. Annu Rev Clin Psychol 9:409–433.

Langner R, Eickhoff SB (2013): Sustaining attention to simple tasks: a meta-analytic review of the neural mechanisms of vigilant attention. Psychol Bull 139:870–900.

Laruelle M, Kegeles LS, Abi-Dargham A (2003): Glutamate, dopamine, and schizophrenia: From pathophysiology to treatment. Ann N Y Acad Sci 1003:138–158.

Lemm S, Blankertz B, Dickhaus T, Müller K-R (2011): Introduction to machine learning for brain imaging. NeuroImage 56:387–399.

Lesh TA, Niendam TA, Minzenberg MJ, Carter CS (2011): Cognitive control deficits in schizophrenia: Mechanisms and meaning. Neuropsychopharmacology 36:316–338.

Liu X, Hairston J, Schrier M, Fan J (2011): Common and distinct networks underlying reward valence and processing stages: a meta-analysis of functional neuroimaging studies. Neurosci Biobehav Rev 35:1219–1236.

Long D, Wang J, Xuan M, Gu Q, Xu X, Kong D, Zhang M (2012): Automatic classification of early Parkinson's disease with multi-modal MR imaging. PLoS One 7:e47714.

Mather M (2016): The affective neuroscience of Aging. Annu Rev Psychol 67:213–238.

Meier TB, Desphande AS, Vergun S, Nair VA, Song J, Biswal BB, Meyerand ME, Birn RM, Prabhakaran V (2012): Support vector machine classification and characterization of age-related reorganization of functional brain networks. NeuroImage 60:601–613.

Movement Disorder Society Task Force on Rating Scales for Parkinson's Disease (2003): The Unified Parkinson's Disease Rating Scale (UPDRS): Status and recommendations. Mov Disord 18:738–750.

Narr KL, Leaver AM (2015): Connectome and schizophrenia. Curr Opin Psychiatry 28:229–235.

Nasreddine ZS, Phillips NA, Bédirian V, Charbonneau S, Whitehead V, Collin I, Cummings JL, Chertkow H (2005): The Montreal Cognitive Assessment, MoCA: A brief screening tool for mild cognitive impairment. J Am Geriatr Soc 53:695–699.

New AB, Robin DA, Parkinson AL, Eickhoff CR, Reetz K, Hoffstaedter F, Mathys C, Sudmeyer M, Grefkes C, Larson CR, Ramig LO, Fox PT, Eickhoff SB (2015): The intrinsic resting state voice network in Parkinson's disease. Hum Brain Mapp 36:1951–1962.

Nieoullon A (2002): Dopamine and the regulation of cognition and attention. Prog Neurobiol 67:53–83.

Obeso JA, Rodríguez-Oroz MC, Benitez-Temino B, Blesa FJ, Guridi J, Marin C, Rodriguez M (2008): Functional organization of the basal ganglia: Therapeutic implications for Parkinson's disease. Mov Disord 23(Suppl 3):S548–S559.

Poewe W, Seppi K, Tanner CM, Halliday GM, Brundin P, Volkmann J, Schrag A-E, Lang AE (2017): Parkinson disease. Nat Rev Dis Primer 3:17013.

Poldrack RA (2011): Inferring mental states from neuroimaging data: From reverse inference to large-scale decoding. Neuron 72:692–697.

Poletti M, Enrici I, Bonuccelli U, Adenzato M (2011): Theory of Mind in Parkinson's disease. Behav Brain Res 219:342–350.

Power JD, Barnes KA, Snyder AZ, Schlaggar BL, Petersen SE (2012): Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. Neuroimage 59:2142–2154.

Prodoehl J, Burciu RG, Vaillancourt DE (2014): Resting state functional magnetic resonance imaging in Parkinson's disease. Curr Neurol Neurosci Rep 14:448.

Radua J, Schmidt A, Borgwardt S, Heinz A, Schlagenhauf F, McGuire P, Fusar-Poli P (2015): Ventral striatal activation during reward processing in psychosis: A neurofunctional meta-analysis. JAMA Psychiatry 72:1243–1251.

Rajah MN, D'Esposito M (2005): Region-specific changes in prefrontal function with age: A review of PET and fMRI studies on working and episodic memory. Brain 128:1964–1983.

Rodriguez M, Rodriguez-Sabate C, Morales I, Sanchez A, Sabate M (2015): Parkinson's disease as a result of aging. Aging Cell 14:293–308.

Rottschy C, Langner R, Dogan I, Reetz K, Laird AR, Schulz JB, Fox PT, Eickhoff SB (2012): Modelling neural correlates of working memory: A coordinate-based meta-analysis. NeuroImage 60:830–846.

Rowe JB, Siebner HR (2012): The motor system and its disorders. NeuroImage 61:464–477.

Sabatinelli D, Fortune EE, Li Q, Siddiqui A, Krafft C, Oliver WT, Beck S, Jeffries J (2011): Emotional perception: Meta-analyses of face and natural scene processing. NeuroImage 54:2524–2533.

Sala-Llonch R, Bartrés-Faz D, Junqué C (2015): Reorganization of brain networks in aging: A review of functional connectivity studies. Front Psychol 6:663.

Satterthwaite TD, Elliott MA, Gerraty RT, Ruparel K, Loughead J, Calkins ME, Eickhoff SB, Hakonarson H, Gur RC, Gur RE, Wolf DH (2013): An improved framework for confound regression and filtering for control of motion artifact in the preprocessing of resting-state functional connectivity data. NeuroImage 64:240–256.

Schilbach L, Derntl B, Aleman A, Caspers S, Clos M, Diederen KMJ, Gruber O, Kogler L, Liemburg EJ, Sommer IE, Müller VI, Cieslik EC, Eickhoff SB (2016): Differential patterns of dysconnectivity in mirror neuron and mentalizing networks in Schizophrenia. Schizophr. Bull 1135–1148.

Schilbach L, Müller VI, Hoffstaedter F, Clos M, Goya-Maldonado R, Gruber O, Eickhoff SB (2014): Meta-analytically informed network analysis of resting state fMRI reveals hyperconnectivity in an introspective socio-affective network in Depression. PLoS One 9:

Schnack HG, Kahn RS (2016): Detecting neuroimaging biomarkers for psychiatric disorders: Sample size matters. Front Psychiatry 7:50.

Schultz W, Dayan P, Montague R (1997): A neural substrate of prediction and reward. Science 275:1593–1599.

Seidler RD, Bernard JA, Burutolu TB, Fling BW, Gordon MT, Gwin JT, Kwak Y, Lipps DB (2010): Motor control and aging: Links to age-related brain structural, functional, and biochemical effects. Neurosci Biobehav Rev 34:721–733.

Smith SM, Fox PT, Miller KL, Glahn DC, Fox PM, Mackay CE, Filippini N, Watkins KE, Toro R, Laird AR, Beckmann CF (2009): Correspondence of the brain's functional architecture during activation and rest. Proc Natl Acad Sci U S A 106:13040–13045.

Spreng RN, Mar RA, Kim ASN (2009): The common neural basis of autobiographical memory, prospection, navigation, theory of mind, and the default mode: a quantitative meta-analysis. J Cogn Neurosci 21:489–510.

Su L, Wang L, Shen H, Feng G, Hu D (2013): Discriminative analysis of non-linear brain connectivity in Schizophrenia: An fMRI study. Front Hum Neurosci 7:

Takahashi H, Koeda M, Oda K, Matsuda T, Matsushima E, Matsuura M, Asai K, Okubo Y (2004): An fMRI study of differential neural response to affective pictures in Schizophrenia. NeuroImage 22:1247–1254.

Tang Y, Wang L, Cao F, Tan L (2012): Identify Schizophrenia using resting-state functional connectivity: An exploratory research and analysis. Biomed Eng Online 11:50.

Tessitore A, Giordano A, De Micco R, Russo A, Tedeschi G (2014): Sensorimotor connectivity in Parkinson's disease: The role of functional neuroimaging. Front Neurol 5:

Toda M, Abi-Dargham A (2007): Dopamine hypothesis of schizophrenia: Making sense of it all. Curr Psychiatry Rep 9:329–336.

Tomlinson CL, Stowe R, Patel S, Rick C, Gray R, Clarke CE (2010): Systematic review of levodopa dose equivalency reporting in Parkinson's disease. Mov Disord off J Mov Disord Soc 25:2649–2653.

van Os J, Kapur S (2009): Schizophrenia. Lancet Lond Engl 374: 635–645.

Varikuti DP, Hoffstaedter F, Genon S, Schwender H, Reid AT, Eickhoff SB (2016): Resting-state test-retest reliability of a priori defined canonical networks over different preprocessing steps. Brain Struct Funct 222:1447–1468.

Varoquaux G, Raamana PR, Engemann DA, Hoyos-Idrobo A, Schwartz Y, Thirion B (2016): Assessing and tuning brain decoders: Cross-validation, caveats, and guidelines. Neuro-Image 145:166–179.

Vergun S, Deshpande AS, Meier TB, Song J, Tudorascu DL, Nair VA, Singh V, Biswal BB, Meyerand ME, Birn RM, Prabhakaran V (2013): Characterizing functional connectivity differences in aging adults using machine learning on resting state fMRI data. Front Comput Neurosci 7:38.

Witt ST, Meyerand ME, Laird AR (2008): Functional neuroimaging correlates of finger tapping task variations: An ALE meta-analysis. NeuroImage 42:343–356.

Wolfers T, Buitelaar JK, Beckmann CF, Franke B, Marquand AF (2015): From estimating activation locality to predicting disorder: A review of pattern recognition for neuroimaging-based psychiatric diagnostics. Neurosci Biobehav Rev 57: 328–349.

Xia M, Wang J, He Y (2013): BrainNet Viewer: A network visualization tool for human brain connectomics. PLoS One 8:e68910.

Yu Y, Shen H, Zhang H, Zeng L-L, Xue Z, Hu D (2013): Functional connectivity-based signatures of Schizophrenia revealed by multiclass pattern analysis of resting-state fMRI from schizophrenic patients and their healthy siblings. Biomed Eng Online 12:10.

Supporting Information

On the integrity of functional brain networks in schizophrenia, Parkinson's disease, and advanced age: Evidence from connectivity-based single-subject classification

Table SI: Functional magnetic resonance imaging parameters

Acquisition Site	Measurement Parameters: Scanner/volumes/TR/TE/FA/voxel size
Schizophrenia Sample	
Mind Research Network, Center for Biomedical Research Excellence (COBRE), The University of New Mexico, Albuquerque, NM, USA	3 T/300/2000/29/75°/3 x 3 x 4 mm^3
University Hospital Göttingen, Germany	3 T/156/2000/30/70°/3 x 3 x 3 mm^3
Parkinson's Disease Sample	
RWTH, University Hospital Aachen, Germany	3 T/165/2200/30/90°/3.1 x 3.1 x 3.1 mm^3
HHU, University Hospital Düsseldorf, Germany	3 T/300/2200/30/90°/3.1 x 3.1 x 3.1 mm^3
Normal Aging Sample	
Research Centre Jülich, Germany	3 T/200/2200/30/80°/3.1 x 3.1 x 3.1 mm^3
1000BRAINS [Caspers et al., 2014], Research Centre Jülich, Germany	3 T/300/2200/30/90°/3.1 x 3.1 x 3.1 mm^3

Measurement parameters:

Scanner: magnetic field strength of the scanner/ number of acquired volumes/TR: repetition time (in ms)/TE: echo time (in ms)/ FA: flip angle/ voxel size.

Table SII: Network coordinates and corresponding brain regions

			Emotional Scene / Face Processing (EmoSF)
			[Sabatinelli et al., 2011]
x	**y**	**z**	**Macroanatomical Region**
4	47	7	R Anterior Cingulate Cortex
42	25	3	R Inferior Frontal Gyrus (p. Triangularis)
-42	25	3	L Inferior Frontal Gyrus (p. Triangularis)
48	17	29	R Inferior Frontal Gyrus (p. Opercularis)
-42	13	27	L Inferior Frontal Gyrus (p. Triangularis)
-2	8	59	L Posterior Medial Frontal
20	-4	-15	R Amygdala
-20	-6	-15	L Amygdala
-20	-33	-4	L Hippocampus
14	-33	-7	R Lingual Gyrus
53	-50	4	R Middle Temporal Gyrus
38	-55	-20	R Anterior Fusiform Gyrus
-40	-55	-22	L Anterior Fusiform Gyrus
38	-76	-16	R Posterior Fusiform Gyrus
-40	-78	-21	L Cerebellum
-4	52	31	L Superior Medial Gyrus
36	25	-3	R Anterior Insula
-38	25	-8	L Inferior Frontal Gyrus (p. Orbitalis)
2	19	25	R Anterior Cingulate Cortex
0	-15	10	Thalamus
-2	-31	-7	Superior Colliculus
-28	-70	-14	L Fusiform Gyrus
46	-68	-4	R Inferior Temporal Gyrus
-48	-72	-4	L Inferior Occipital Gyrus

			Cognitive Emotion Regulation (ER)
			[Buhle et al., 2014]
x	**y**	**z**	**Macroanatomical Region**
48	24	9	R Inferior Frontal Gyrus (p. Triangularis)

x	y	z	Macroanatomical Region
42	21	45	R Middle Frontal Gyrus
9	30	39	R Superior Medial Gyrus
0	-9	63	L Posterior Medial Frontal
-3	24	30	L Anterior Cingulate Cortex
-33	3	54	L Middle Frontal Gyrus
-36	21	-3	L Anterior Insula
-42	45	-6	L Inferior Frontal Gyrus (p. Orbitalis)
63	-51	39	R Inferior Parietal Lobule
-42	-66	42	L Angular Gyrus
-63	-51	-21	L Inferior Temporal Gyrus
-51	-39	3	L Middle Temporal Gyrus
30	-3	-15	R Amygdala
-18	-3	-15	L Amygdala

Theory-of-Mind Cognition (ToM)
[Bzdok et al., 2012]

x	y	z	Macroanatomical Region
0	52	-12	R Mid Orbital Gyrus
2	58	12	R Superior Medial Gyrus
-8	56	30	L Superior Medial Gyrus
2	-56	30	L Precuneus
56	-50	18	R Superior Temporal Gyrus
-48	-56	24	L Angular Gyrus
54	-2	-20	R Anterior Middle Temporal Gyrus
-54	-2	-24	L Anterior Middle Temporal Gyrus
52	-18	-12	R Middle Temporal Gyrus
-54	-28	-4	L Middle Temporal Gyrus
50	-34	0	R Posterior Superior Temporal Sulcus
-58	-44	4	L Posterior Superior Temporal Sulcus
54	28	6	R Inferior Frontal Gyrus (p. Triangularis)
-48	30	-12	L Inferior Frontal Gyrus (p. Orbitalis)
48	-72	8	R Occipital Lobe (V5/MT)

Empathic Processing (Empathy)
[Bzdok et al., 2012]

x	y	z	Macroanatomical Region
2	56	18	L Superior Medial Gyrus
36	22	-8	R Inferior Frontal Gyrus (p. Orbitalis)

x	y	z	Macroanatomical Region
-30	20	4	L Anterior Insula
50	12	-8	R Anterior Insula
-44	24	-6	L Inferior Frontal Gyrus (p. Orbitalis)
-4	18	50	L Posterior Medial Frontal
-2	28	20	L Anterior Cingulate Cortex
-4	42	18	L Anterior Cingulate Cortex
-2	-32	28	Posterior Cingulate Cortex
52	-58	22	R Posterior Superior Temporal Gyrus
-56	-58	22	L Posterior Superior Temporal Gyrus
22	-2	-16	R Amygdala
54	-8	-16	R Middle Temporal Gyrus
52	-36	2	R Posterior Superior Temporal Sulcus
-12	-4	12	L Anterior Thalamus
6	-32	2	R Posterior Thalamus
26	-26	-12	R Hippocampus
2	-20	-12	Midbrain
14	4	0	R Globus Pallidum

Reward-related Decision Making (Rew)
[Liu et al., 2011]

x	y	z	Macroanatomical Region
12	10	-6	R Nucleus Caudate
-10	8	-4	L Pallidum
36	20	-6	R Anterior Insula
-32	20	-4	L Anterior Insula
0	24	40	L Superior Medial Gyrus
0	54	-8	L Mid Orbital Gyrus
24	-2	-16	R Amygdala
6	-14	8	R Thalamus
-6	-16	8	L Thalamus
0	8	48	L Posterior Medial Frontal Gyrus
8	-18	-10	R Brainstem
-6	-18	-10	L Brainstem
2	44	20	L Anterior Cingulate Cortex
-24	2	52	L Middle Frontal Gyrus
-38	-4	6	L Insula
24	40	-14	R Superior Orbital Gyrus
-16	42	-14	L Superior Orbital Gyrus
40	32	32	R Middle Frontal Gyrus

x	y	z	Macroanatomical Region
-28	-56	48	L Inferior Parietal Lobule
28	-58	50	R Superior Parietal Lobule
0	-32	32	L Posterior Cingulate Cortex
-36	50	10	L Middle Frontal Gyrus
-46	42	-4	L Inferior Frontal Gyrus (p. Orbitalis)
30	4	50	R Middle Frontal Gyrus
-22	30	48	L Superior Frontal Gyrus

Autobiographical Memory (AM)
[Spreng et al., 2009]

x	y	z	Macroanatomical Region
-1	-53	21	L Precuneus
-26	-28	-17	L Parahippocampal Gyrus
-49	-61	31	L Angular Gyrus
-2	51	-11	L Mid Orbital Gyrus
-60	-9	-18	L Middle Temporal Gyrus
-50	27	-12	L Inferior Frontal Gyrus (p. Orbitalis)
26	-33	-15	R Fusiform Gyrus
-1	20	57	L Posterior Medial Frontal
55	-58	30	R Angular Gyrus
-47	9	46	L Precentral Gyrus
-42	53	7	L Middle Frontal Gyrus
26	-14	-23	R Parahippocampal Gyrus
54	-5	-20	R Middle Temporal Gyrus
-39	13	-41	L Inferior Temporal Gyrus
-38	-82	38	L Middle Occipital Gyrus
-48	29	17	L Inferior Frontal Gyrus (p. Triangularis)
-11	62	9	L Superior Medial Gyrus
4	-8	2	Thalamus
-4	39	16	L Anterior Cingulate Cortex
-5	-34	36	L Midcingulate Cortex
-29	16	51	L Middle Frontal Gyrus
31	1	-26	R Amygdala

Semantic Memory (SM)
[Binder et al., 2009]

x	y	z	Macroanatomical Region
-46	-69	28	L Angular Gyrus

x	y	z	Macroanatomical Region
-50	-56	31	L Angular Gyrus
-64	-44	-4	L Posterior Middle Temporal Gyrus
-47	-24	-17	L Middle Temporal Gyrus
-40	-12	-30	L Inferior Temporal Gyrus
-8	-57	17	L Precuneus
-20	36	44	L Superior Frontal Gyrus
-53	27	-4	L Inferior Frontal Gyrus (p. Orbitalis)
54	-59	30	R Angular Gyrus
43	-72	31	R Middle Occipital Gyrus
-1	51	-7	L Mid Orbital Gyrus
-5	56	24	L Superior Medial Gyrus
-31	-34	-16	L Fusiform Gyrus
-8	29	-10	L Anterior Cingulate Cortex
-46	25	23	L Inferior Frontal Gyrus (p. Triangularis)
64	-41	-2	R Posterior Middle Temporal Gyrus
-43	-53	55	L Inferior Parietal Lobule
-1	-18	40	L Midcingulate Cortex
-2	-56	46	L Precuneus
51	20	26	R Inferior Frontal Gyrus (p. Triangularis)
64	-38	32	R Supramarginal Gyrus
-23	26	-16	L Inferior Frontal Gyrus (p. Orbitalis)
-5	-39	40	L Midcingulate Cortex

Working Memory (WM)
[Rottschy et al., 2012]

x	y	z	Macroanatomical Region
-32	22	-2	L Anterior Insula
-48	10	26	L Inferior Frontal Gyrus (p. Opercularis)
-46	26	24	L Inferior Frontal Gyrus (p. Triangularis)
-38	50	10	L Anterior Middle Frontal Gyrus
36	22	-6	R Anterior Insula
50	14	24	R Inferior Frontal Gyrus (p. Triangularis)
44	34	32	R Middle Frontal Gyrus
38	54	6	R Anterior Middle Frontal Gyrus

x	y	z	Macroanatomical Region
2	18	48	L Posterior Medial Frontal
-28	0	56	L Posterior Middle Frontal Gyrus
30	2	56	R Posterior Middle Frontal Gyrus
-42	-42	46	L Inferior Parietal Lobule/Intraparietal Sulcus
-34	-52	48	L Inferior Parietal Lobule/Intraparietal Sulcus
-24	-66	54	L Superior Parietal Lobule
42	-44	44	R Inferior Parietal Lobule/Intraparietal Sulcus
32	-58	48	R Angular Gyrus/Intraparietal Sulcus
16	-66	56	R Superior Parietal Lobule
-12	-12	12	L Thalamus
-16	2	14	L Nucleus Caudate
-16	0	2	L Globus Pallidum
12	-10	10	R Thalamus
-34	-66	-20	L Cerebelum/Fusiform Gyrus
32	-64	-18	R Cerebelum/Fusiform Gyrus

Cognitive Action Control (CogAC)
[Cieslik et al., 2015]

x	y	z	Macroanatomical Region
36	22	-4	R Anterior Insula
2	16	48	L Posterior Medial Frontal
48	12	30	R Inferior Frontal Gyrus (p. Opercularis)
36	2	54	R Middle Frontal Gyrus
48	30	24	R Inferior Frontal Gyrus (p. Triangularis)
-38	-44	46	L Inferior Parietal Lobule/Intraparietal Sulcus
-24	-66	48	L Superior Parietal Lobule
40	-46	46	R Inferior Parietal Lobule/Intraparietal Sulcus
60	-44	24	R Supramarginal Gyrus
30	-62	52	R Superior Parietal Lobule
-44	10	30	L Precentral Gyrus
-34	20	-4	L Anterior Insula
-26	2	52	L Middle Frontal Gyrus

x	y	z	Macroanatomical Region
6	-18	-2	R Thalamus
-40	-66	-10	L Inferior Occipital Gyrus
48	19	6	R Inferior Frontal Gyrus (p. Opercularis)
8	29	30	R Midcingulate Cortex
-45	27	30	L Inferior Frontal Gyrus (p. Triangularis)
11	7	7	R Nucleus Caudate

Vigilant Attention (VigAtt)
[Langner and Eickhoff, 2013]

x	y	z	Macroanatomical Region
-2	8	50	L Posterior Medial Frontal
8	32	46	R Superior Medial Gyrus
0	26	34	L Midcingulate Cortex
50	8	32	R Precentral Gyrus
40	22	-4	R Anterior Insula
46	36	20	R Anterior Middle Frontal Gyrus
-40	-12	60	L Precentral Gyrus
-46	-68	-6	L Inferior Occipital Gyrus
-48	8	30	L Precentral Gyrus
62	-38	17	R Inferior Parietal Lobe
8	-12	6	R Thalamus
32	-90	4	R Middle Occipital Gyrus
-42	12	-2	L Anterior Insula
-10	-14	6	L Thalamus
6	-58	-18	Cerebellar Vermis
44	-44	46	R Inferior Parietal Lobule

Mirror Neuron System (MNS)
[Caspers et al., 2010]

x	y	z	Macroanatomical Region
-56	8	28	L Precentral Gyrus
-54	6	40	L Precentral Gyrus
58	16	10	R Inferior Frontal Gyrus (p. Opercularis)
44	-54	-20	R Fusiform Gyrus
-38	-40	50	L Inferior Parietal Lobule/Intraparietal Sulcus

x	y	z	Macroanatomical Region
51	-36	50	R Inferior Parietal Lobule/Intraparietal Sulcus
-1	16	52	L Posterior Medial Frontal
-54	-50	10	L Posterior Middle Temporal Gyrus
-52	-70	6	L Occipital Lobe (V5)
54	-64	4	R Occipital Lobe (V5)
30	-62	63	R Superior Parietal Lobule

Motor Execution (Motor)
[Witt et al., 2008]

x	y	z	Macroanatomical Region
-39	-21	54	L Postcentral Gyrus
41	-16	57	R Precentral Gyrus
-3	-2	54	L Posterior Medial Frontal
-57	2	32	L Precentral Gyrus
-53	-24	21	L Supramarginal Gyrus
45	-38	48	R Inferior Parietal Lobule
-23	-7	1	L Globus Pallidum
25	-8	3	R Globus Pallidum
-22	-52	26	L Cerebellum
18	-54	-22	R Cerebellum

R= right; **L** = left; for consistency coordinates (MNI-space) are assigned to the most probable brain areas as revealed by the SPM Anatomy Toolbox (Version 2.1) [Eickhoff et al., 2005; Eickhoff et al., 2006; Eickhoff et al., 2007].

Table III: Classification results of the support vector machine of all groups based on specific networks (balanced accuracy)

Network (abbr.)	SCZ vs. HC$_{SCZ}$	PD vs. HC$_{PD}$	Old vs. Young
EmoSF	73%	64%	88%
ER	71%	69%	78%
ToM	60%	67%	79%
Empathy	71%	63%	78%
Rew	68%	67%	87%
AM	62%	76%	80%
SM	61%	69%	84%
WM	63%	66%	79%
CogAC	68%	62%	74%
VigAtt	68%	66%	81%
MNS	64%	58%	84%
Motor	61%	71%	80%

Balanced accuracy is calculated as the average proportion of subjects correctly classified as patients (PD, SCZ) or advanced age versus healthy or younger age, respectively.

Table SIV: Classification results of the support vector machine of all groups based on specific networks (d')

Network (abbr.)	SCZ vs. HC$_{SCZ}$	PD vs. HC$_{PD}$	Old vs. Young
EmoSF	1.19	0.69	2.33
ER	1.13	1.00	1.52
ToM	0.57	0.89	1.56
Empathy	1.12	0.68	1.52
Rew	0.92	0.86	2.23
AM	0.63	1.36	1.67
SM	0.54	0.99	1.99
WM	0.63	0.79	1.59
CogAC	0.95	0.60	1.25
VigAtt	0.92	0.79	1.71
MNS	0.72	0.38	1.95
Motor	0.59	1.06	1.67

d': sensitivity index calculated as z (true positive rate) – z (false positive rate).

Functional Networks

AM - Autobiographical memory
Motor - Motor execution
SM - Semantic memory
ToM - Theory-of-mind cognition
WM - Working memory
Rew - Reward-related decision making
VigAtt - Vigilant attention
ER - Cognitive emotion regulation
MNS - Mirror neuron system
CogAC - Cognitive action control
Empathy – Empathic processing
EmoSF - Emotional scene / face processing
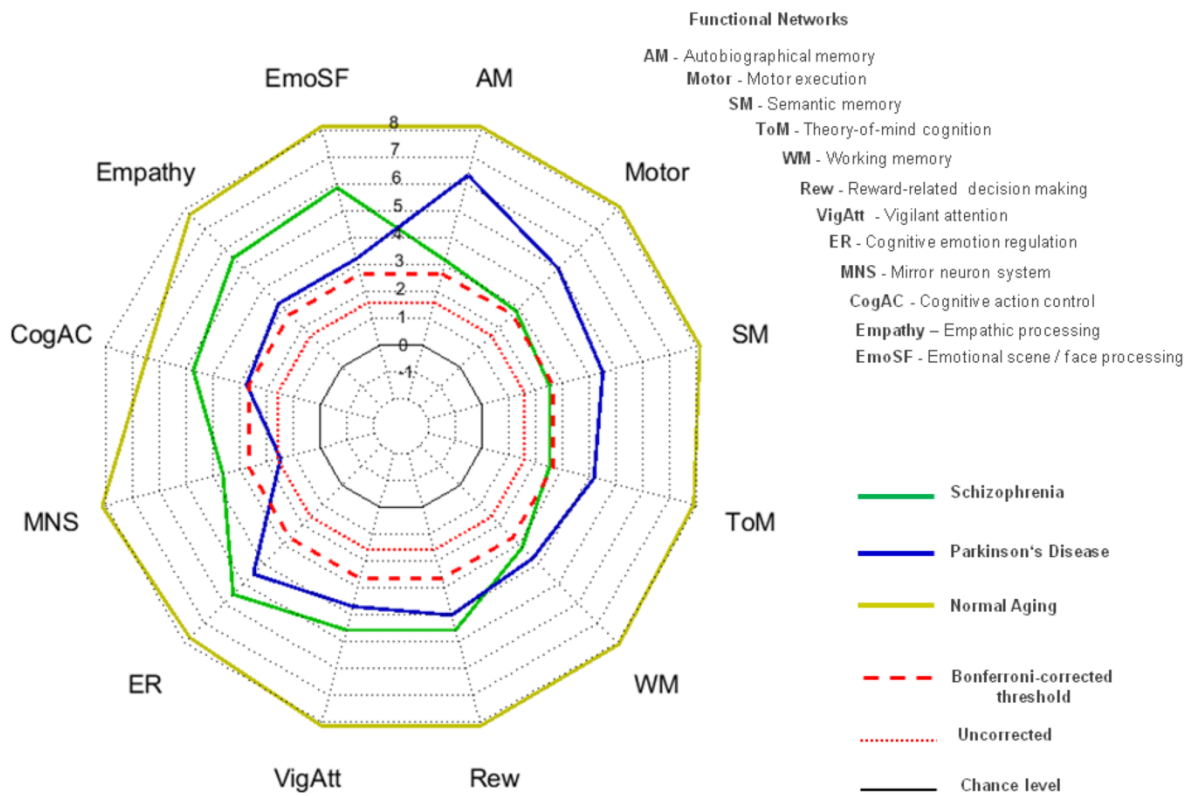
Figure S1: Group classification results of the support vector machine (z-values)

Polar plot of z-standardized accuracies (corrected for multiple comparisons) of group classification based on all 12 networks for schizophrenia (in green), Parkinson's disease (in blue) and normal aging (in yellow).
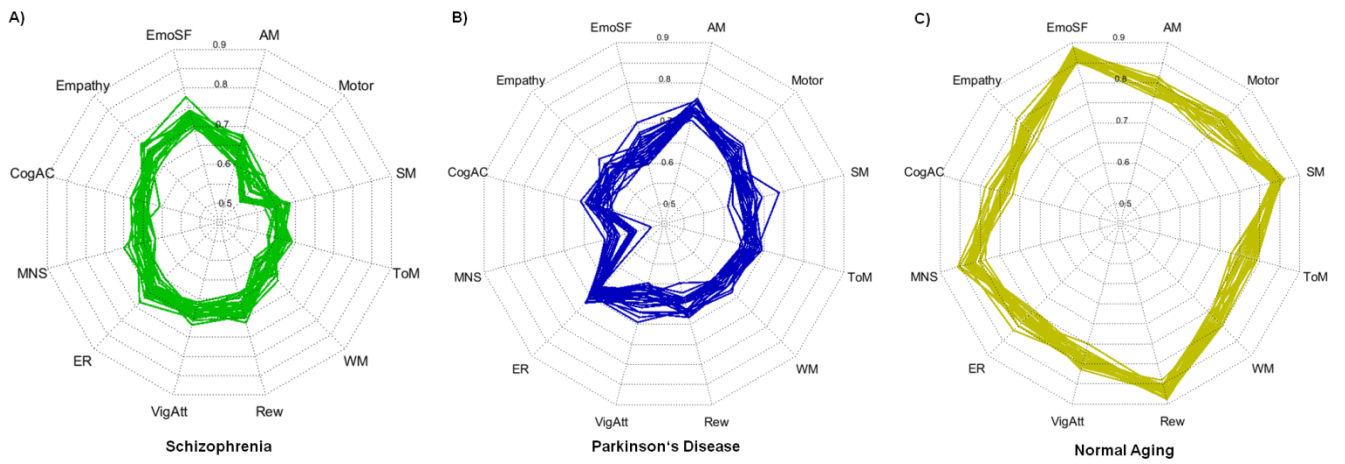
Figure S2: Variance of group classification results of the support vector machine (accuracies)

Polar plot of variance for group classification accuracies over all 25 repetitions in the outer loop based on all 12 networks for **A)** schizophrenia (in green), **B)** Parkinson's disease (in blue) and **C)** normal aging (in yellow).

Table SV: Differences in classification performance between networks within schizophrenia

Network (abbr.) comparison	Mean difference (Acc.)	T	*P*
EmoSF - ER	**3.320**	**6.211**	**< 0.0001**
EmoSF - ToM	**11.240**	**23.735**	**< 0.0001**
EmoSF - Empathy	2.920	6.704	< 0.0001
EmoSF - Rew	**5.080**	**10.098**	**< 0.0001**
EmoSF - AM	**8.720**	**14.498**	**< 0.0001**
EmoSF - SM	**11.840**	**28.255**	**< 0.0001**
EmoSF - WM	**10.800**	**22.358**	**< 0.0001**
EmoSF - Cog	7.040	14.735	< 0.0001
EmoSF - VigAtt	**4.360**	**9.375**	**< 0.0001**
EmoSF - MNS	**7.280**	**14.510**	**< 0.0001**
EmoSF - Motor	**15.640**	**22.506**	**< 0.0001**
ER - ToM	7.920	15.540	< 0.0001
ER - Empathy	**0.400**	**0.679**	**0.5034**
ER - Rew	1.760	2.903	0.0078
ER - AM	5.400	7.991	< 0.0001
ER - SM	8.520	12.649	< 0.0001
ER - WM	7.480	11.821	< 0.0001
ER - Cog	**3.720**	**7.318**	**< 0.0001**
ER - VigAtt	1.040	1.996	0.05745
ER - MNS	3.960	6.678	< 0.0001
ER - Motor	12.320	16.991	< 0.0001
ToM - Empathy	**8.320**	**17.512**	**< 0.0001**
ToM - Rew	6.160	11.083	< 0.0001
ToM - AM	2.520	4.016	0.0005
ToM - SM	0.600	1.200	0.2419
ToM - WM	0.440	0.751	0.4602
ToM - Cog	**4.200**	**9.635**	**< 0.0001**
ToM - VigAtt	6.880	14.633	< 0.0001
ToM - MNS	3.960	7.251	< 0.0001
ToM - Motor	4.400	7.738	< 0.0001
Empathy - Rew	**2.160**	**4.018**	**0.0005**
Empathy - AM	**5.800**	**10.307**	**< 0.0001**
Empathy - SM	**8.920**	**15.942**	**< 0.0001**
Empathy - WM	**7.880**	**13.114**	**< 0.0001**
Empathy - Cog	4.120	7.771	< 0.0001
Empathy - VigAtt	**1.440**	**3.490**	**0.0019**

Empathy - MNS	**4.360**	**7.966**	**< 0.0001**
Empathy - Motor	**12.720**	**24.253**	**< 0.0001**
Rew - AM	3.640	7.888	< 0.0001
Rew - SM	6.760	12.987	< 0.0001
Rew - WM	5.720	9.981	< 0.0001
Rew - Cog	**1.960**	**4.876**	**< 0.0001**
Rew - VigAtt	0.720	1.705	0.1012
Rew - MNS	2.200	3.755	0.0010
Rew - Motor	10.560	14.881	< 0.0001
AM - SM	3.120	6.227	< 0.0001
AM - WM	2.080	3.125	0.0046
AM - Cog	**1.680**	**3.072**	**0.0052**
AM - VigAtt	4.360	8.557	< 0.0001
AM - MNS	1.440	2.138	0.0429
AM - Motor	6.920	9.538	< 0.0001
SM - WM	1.040	2.279	0.0319
SM - Cog	**4.800**	**11.314**	**< 0.0001**
SM - VigAtt	7.480	13.647	< 0.0001
SM - MNS	4.560	8.771	< 0.0001
SM - Motor	3.800	4.961	< 0.0001
WM - Cog	**3.760**	**7.556**	**< 0.0001**
WM - VigAtt	6.440	10.398	< 0.0001
WM - MNS	3.520	7.179	< 0.0001
WM - Motor	4.840	5.933	< 0.0001
Cog - VigAtt	**2.680**	**6.650**	**< 0.0001**
Cog - MNS	**0.240**	**0.448**	**0.6585**
Cog - Motor	**8.600**	**12.820**	**< 0.0001**
VigAtt - MNS	2.920	4.985	< 0.0001
VigAtt - Motor	11.280	17.871	< 0.0001
MNS - Motor	8.360	13.116	< 0.0001

Comparison between networks with highest classification performance and all other networks in schizophrenia (in bold); significance threshold $P_{corr} < 0.001$.

Table SVI: Differences in classification performance between networks within Parkinson's disease

Network (abbr.) comparison	Mean difference (Acc.)	T	*P*
EmoSF - ER	6.000	8.721	< 0.0001

EmoSF - ToM	**3.560**	**4.799**	**< 0.0001**
EmoSF - Empathy	0.680	1.034	0.3114
EmoSF - Rew	1.720	3.149	0.0043
EmoSF - AM	**10.400**	**18.196**	**< 0.0001**
EmoSF - SM	**3.080**	**4.447**	**0.0002**
EmoSF - WM	1.520	2.354	0.0271
EmoSF - Cog	1.480	2.128	0.0438
EmoSF - VigAtt	0.120	0.197	0.8455
EmoSF - MNS	8.200	10.645	< 0.0001
EmoSF - Motor	**5.440**	**7.566**	**< 0.0001**
ER - ToM	**2.440**	**5.160**	**< 0.0001**
ER - Empathy	6.680	11.148	< 0.0001
ER - Rew	4.280	7.546	< 0.0001
ER - AM	**4.400**	**10.184**	**< 0.0001**
ER - SM	**2.920**	**4.276**	**0.0003**
ER - WM	4.480	10.267	< 0.0001
ER - Cog	7.480	13.724	< 0.0001
ER - VigAtt	6.120	10.185	< 0.0001
ER - MNS	14.200	21.168	< 0.0001
ER - Motor	**0.560**	**1.212**	**0.2374**
ToM - Empathy	**4.240**	**9.021**	**< 0.0001**
ToM - Rew	**1.840**	**3.100**	**0.0049**
ToM - AM	6.840	15.085	< 0.0001
ToM - SM	0.480	0.762	0.4536
ToM - WM	**2.040**	**4.270**	**0.0003**
ToM - Cog	**5.040**	**10.325**	**< 0.0001**
ToM - VigAtt	**3.680**	**5.075**	**< 0.0001**
ToM - MNS	**11.760**	**16.039**	**< 0.0001**
ToM - Motor	1.880	3.969	0.0006
Empathy - Rew	2.400	3.811	0.0008
Empathy - AM	**11.080**	**29.641**	**< 0.0001**
Empathy - SM	**3.760**	**6.317**	**< 0.0001**
Empathy - WM	2.200	4.260	0.0003
Empathy - Cog	0.800	1.469	0.1549
Empathy - VigAtt	0.560	0.854	0.4017
Empathy - MNS	7.520	11.205	< 0.0001
Empathy - Motor	**6.120**	**9.394**	**< 0.0001**
Rew - AM	**8.680**	**18.136**	**< 0.0001**
Rew - SM	**1.360**	**2.134**	**0.0433**
Rew - WM	0.200	0.447	0.6587

Rew - Cog	3.200	5.409	< 0.0001
Rew - VigAtt	1.840	3.145	0.0044
Rew - MNS	9.920	12.435	< 0.0001
Rew - Motor	**3.720**	**7.566**	**< 0.0001**
AM - SM	7.320	12.510	< 0.0001
AM - WM	**8.880**	**25.177**	**< 0.0001**
AM - Cog	**11.880**	**25.859**	**< 0.0001**
AM - VigAtt	**10.520**	**18.683**	**< 0.0001**
AM - MNS	**18.600**	**29.167**	**< 0.0001**
AM - Motor	4.960	11.639	< 0.0001
SM - WM	**1.560**	**3.019**	**0.0059**
SM - Cog	**4.560**	**6.384**	**< 0.0001**
SM - VigAtt	**3.200**	**3.687**	**0.0016**
SM - MNS	**11.280**	**14.155**	**< 0.0001**
SM - Motor	2.360	3.432	0.0022
WM - Cog	3.000	6.000	< 0.0001
WM - VigAtt	1.640	2.605	0.0155
WM - MNS	9.720	14.409	< 0.0001
WM - Motor	**3.920**	**8.699**	**< 0.0001**
Cog - VigAtt	1.360	2.091	0.0473
Cog - MNS	6.720	9.333	< 0.0001
Cog - Motor	**6.920**	**11.815**	**< 0.0001**
VigAtt - MNS	8.080	9.150	< 0.0001
VigAtt - Motor	**5.560**	**7.835**	**< 0.0001**
MNS - Motor	**13.640**	**18.569**	**< 0.0001**

Comparison between networks with highest classification performance and all other networks in Parkinson's disease (in bold); significance threshold $P_{corr} < 0.001$.

Table SVII: Group differences between schizophrenia and Parkinson's disease classification based on specific networks

Network (abbr.)	Mean difference (Acc.)	T	*P*
EmoSF	9.120	14.028	< 0.0001
ER	0.200	0.376	0.7088
ToM	5.680	12.267	< 0.0001
Empathy	6.880	12.561	< 0.0001
Rew	2.320	4.011	0.0002
AM	10.00	19.240	< 0.0001
SM	5.800	8.985	< 0.0001

WM	3.200	6.137	< 0.0001
CogAC	3.560	6.908	< 0.0001
VigAtt	4.880	7.593	< 0.0001
MNS	10.040	14.013	< 0.0001
Motor	11.960	18.011	< 0.0001

Networks with highest classification performance in schizophrenia (in green); networks with highest classification performance in Parkinson's disease (in blue); significance threshold $P_{corr} < 0.001$.

Table SVIII: Differences in classification performance between networks within normal aging

Network (abbr.) comparison	Mean difference (Acc.)	T	P
EmoSF - ER	9.640	24.386	< 0.0001
EmoSF - ToM	10.080	20.307	< 0.0001
EmoSF - Empathy	10.080	29.131	< 0.0001
EmoSF - Rew	0.080	0.267	0.7917
EmoSF - AM	8.000	24.495	< 0.0001
EmoSF - SM	2.400	8.668	< 0.0001
EmoSF - WM	8.840	23.167	< 0.0001
EmoSF - Cog	12.160	29.018	< 0.0001
EmoSF - VigAtt	8.080	22.727	< 0.0001
EmoSF - MNS	4.400	12.882	< 0.0001
EmoSF - Motor	8.400	18.046	< 0.0001
ER - ToM	0.440	0.910	0.3717
ER - Empathy	0.440	1.204	0.2404
ER - Rew	9.560	25.526	< 0.0001
ER - AM	1.640	4.194	0.0003
ER - SM	7.240	19.286	< 0.0001
ER - WM	0.800	1.668	0.1083
ER - Cog	2.520	4.573	0.0001
ER - VigAtt	1.560	3.576	0.0015
ER - MNS	5.240	10.324	< 0.0001
ER - Motor	1.240	2.055	0.0510
ToM - Empathy	0.000	0.000	1.0000
ToM - Rew	10.000	31.109	< 0.0001
ToM - AM	2.080	4.578	0.0001
ToM - SM	7.680	17.663	< 0.0001

ToM - WM	1.240	3.303	0.0030
ToM - Cog	2.080	3.864	0.0007
ToM - VigAtt	2.000	4.082	0.0004
ToM - MNS	5.680	9.435	< 0.0001
ToM - Motor	1.680	2.929	0.0073
Empathy - Rew	10.000	30.151	< 0.0001
Empathy - AM	2.080	5.316	< 0.0001
Empathy - SM	7.680	26.755	< 0.0001
Empathy - WM	1.240	3.059	0.0054
Empathy - Cog	2.080	4.219	0.0003
Empathy - VigAtt	2.000	5.477	< 0.0001
Empathy - MNS	5.680	11.537	< 0.0001
Empathy - Motor	1.680	2.959	0.0068
Rew - AM	7.920	28.654	< 0.0001
Rew - SM	2.320	7.632	< 0.0001
Rew - WM	8.760	27.573	< 0.0001
Rew - Cog	12.080	28.233	< 0.0001
Rew - VigAtt	8.000	21.381	< 0.0001
Rew - MNS	4.320	10.832	< 0.0001
Rew - Motor	8.320	16.671	< 0.0001
AM - SM	5.600	17.421	< 0.0001
AM - WM	0.840	2.227	0.0356
AM - Cog	4.160	9.744	< 0.0001
AM - VigAtt	0.080	0.219	0.8283
AM - MNS	3.600	8.647	< 0.0001
AM - Motor	0.400	0.747	0.4623
SM - WM	6.440	21.435	< 0.0001
SM - Cog	9.760	23.380	< 0.0001
SM - VigAtt	5.680	16.235	< 0.0001
SM - MNS	2.000	5.410	< 0.0001
SM - Motor	6.000	14.412	< 0.0001
WM - Cog	3.320	6.565	< 0.0001
WM - VigAtt	0.760	2.156	0.0413
WM - MNS	4.440	9.320	< 0.0001
WM - Motor	0.440	0.938	0.3578
Cog - VigAtt	4.080	7.955	< 0.0001
Cog - MNS	7.760	15.809	< 0.0001
Cog - Motor	3.760	6.014	< 0.0001
VigAtt - MNS	3.680	9.227	< 0.0001
VigAtt - Motor	0.320	0.730	0.4727

MNS - Motor	4.000	8.281	< 0.0001

Significance threshold $P_{corr} < 0.0001$.

Table SIX: Group differences between schizophrenia and normal aging classification based on specific networks

Network (abbr.)	Mean difference (Acc.)	T	*P*
EmoSF	14.480	35.469	< 0.0001
ER	8.160	14.995	< 0.0001
ToM	15.640	31.139	< 0.0001
Empathy	7.320	16.128	< 0.0001
Rew	19.480	43.003	< 0.0001
AM	15.200	29.173	< 0.0001
SM	23.920	57.946	< 0.0001
WM	16.440	32.283	< 0.0001
CogAC	9.360	18.938	< 0.0001
VigAtt	10.760	24.556	< 0.0001
MNS	17.360	33.451	< 0.0001
Motor	21.720	31.016	< 0.0001

Significance threshold $P_{corr} < 0.001$.

Table SX: Group differences between Parkinson's disease and normal aging classification based on specific networks

Network (abbr.)	Mean difference (Acc.)	T	*P*
EmoSF	23.600	40.347	< 0.0001
ER	7.960	16.335	< 0.0001
ToM	9.960	19.051	< 0.0001
Empathy	14.200	27.574	< 0.0001
Rew	21.800	47.139	< 0.0001
AM	5.200	14.720	< 0.0001
SM	18.120	32.350	< 0.0001
WM	13.240	31.323	< 0.0001
CogAC	12.920	22.726	< 0.0001
VigAtt	15.640	25.407	< 0.0001
MNS	27.400	40.190	< 0.0001
Motor	9.760	18.737	< 0.0001

Significance threshold $P_{corr} < 0.001$.

References

Binder JR, Desai RH, Graves WW, Conant LL (2009): Where Is the Semantic System? A Critical Review and Meta-Analysis of 120 Functional Neuroimaging Studies. Cereb Cortex 19:2767–2796.

Buhle JT, Silvers JA, Wager TD, Lopez R, Onyemekwu C, Kober H, Weber J, Ochsner KN (2014): Cognitive reappraisal of emotion: a meta-analysis of human neuroimaging studies. Cereb Cortex 24:2981–2990.

Bzdok D, Schilbach L, Vogeley K, Schneider K, Laird AR, Langner R, Eickhoff SB (2012): Parsing the neural correlates of moral cognition: ALE meta-analysis on morality, theory of mind, and empathy. Brain Struct Funct 217:783–796.

Caspers S, Moebus S, Lux S, Pundt N, Schütz H, Mühleisen TW, Gras V, Eickhoff SB, Romanzetti S, Stöcker T, Stirnberg R, Kirlangic ME, Minnerop M, Pieperhoff P, Mödder U, Das S, Evans AC, Jöckel K-H, Erbel R, Cichon S, Nöthen MM, Sturma D, Bauer A, Jon Shah N, Zilles K, Amunts K (2014): Studying variability in human brain aging in a population-based German cohort—rationale and design of 1000BRAINS. Front Aging Neurosci 6.

Caspers S, Zilles K, Laird AR, Eickhoff SB (2010): ALE meta-analysis of action observation and imitation in the human brain. NeuroImage 50:1148–1167.

Cieslik EC, Mueller VI, Eickhoff CR, Langner R, Eickhoff SB (2015): Three key regions for supervisory attentional control: evidence from neuroimaging meta-analyses. Neurosci Biobehav Rev 48:22–34.

Eickhoff SB, Heim S, Zilles K, Amunts K (2006): Testing anatomically specified hypotheses in functional imaging using cytoarchitectonic maps. NeuroImage 32:570–582.

Eickhoff SB, Paus T, Caspers S, Grosbras M-H, Evans AC, Zilles K, Amunts K (2007): Assignment of functional activations to probabilistic cytoarchitectonic areas revisited. NeuroImage 36:511–521.

Eickhoff SB, Stephan KE, Mohlberg H, Grefkes C, Fink GR, Amunts K, Zilles K (2005): A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. NeuroImage 25:1325–1335.

Langner R, Eickhoff SB (2013): Sustaining attention to simple tasks: a meta-analytic review of the neural mechanisms of vigilant attention. Psychol Bull 139:870–900.

Liu X, Hairston J, Schrier M, Fan J (2011): Common and distinct networks underlying reward valence and processing stages: a meta-analysis of functional neuroimaging studies. Neurosci Biobehav Rev 35:1219–1236.

Rottschy C, Langner R, Dogan I, Reetz K, Laird AR, Schulz JB, Fox PT, Eickhoff SB (2012): Modelling neural correlates of working memory: a coordinate-based meta-analysis. NeuroImage 60:830–846.

Sabatinelli D, Fortune EE, Li Q, Siddiqui A, Krafft C, Oliver WT, Beck S, Jeffries J (2011): Emotional perception: meta-analyses of face and natural scene processing. NeuroImage 54:2524–2533.

Spreng RN, Mar RA, Kim ASN (2009): The common neural basis of autobiographical memory, prospection, navigation, theory of mind, and the default mode: a quantitative meta-analysis. J Cogn Neurosci 21:489–510.

Witt ST, Meyerand ME, Laird AR (2008): Functional neuroimaging correlates of finger tapping task variations: An ALE meta-analysis. NeuroImage 42:343–356.

STUDY 3

Predicting Personality from Network-based Resting-State Functional Connectivity

Alessandra D. Nostro[1,2,3], Veronika I. Müller[1,2,3], Deepthi P. Varikuti[1,2,3], Rachel N. Pläschke[1,2,3], Felix Hoffstaedter[2,3], Robert Langner[1,2,3], Kaustubh R. Patil[1,3] and Simon B. Eickhoff [1,2,3]

[1]Institute of Systems Neuroscience, Medical Faculty, Heinrich-Heine University Düsseldorf, Universitätstraße 1; 40225 Düsseldorf, Germany
[2]Institute of Clinical Neuroscience and Medical Psychology, Heinrich-Heine University Düsseldorf, Universitätstraße 1; 40225 Düsseldorf, Germany
[3]Institute of Neuroscience and Medicine (INM-1,7), Research Centre Jülich, Wilhelm-Johnen-Straße; 52425 Jülich, Germany

Own contributions   Data processing
                     Critically revising the paper
                     Total contribution: 15%

Abstract

Personality is associated with variation in all kinds of mental faculties, including affective, social, executive and memory functioning. The intrinsic dynamics of neural networks underlying these mental functions are reflected in their functional connectivity at rest (RSFC). We therefore aimed to probe whether connectivity in functional networks allow predicting individual scores of the five-factor personality model and potential gender differences thereof.

We assessed nine meta-analytically derived functional networks, representing social, affective, executive and mnemonic systems. RSFC of all networks was computed in a sample of 210 males and 210 well-matched females and in a replication sample of 155 males and 155 females. Personality scores were predicted using relevance vector machine in both samples. Cross-validation prediction accuracy was defined as the correlation between true and predicted scores.

RSFC within networks representing social, affective, mnemonic and executive systems significantly predicted self-reported levels of Extraversion, Neuroticism, Agreeableness and Openness. RSFC patterns of most networks, however, predicted personality traits only either in males or in females.

Personality traits can be predicted by patterns of RSFC in specific functional brain networks, providing new insights into the neurobiology of personality. However, as most associations were gender-specific, RSFC–personality relations should not be considered independently of gender.

1.   Introduction

Interindividual differences in personality permeate all aspects of life, from affective and cognitive functioning to social relationships. One of the most comprehensive and most widely recognized models of personality is the Five Factor Model (FFM; Costa & McCrae, 1992), consisting of five broad dimensions: Openness to experience/Intellect, Extraversion, Neuroticism, Agreeableness, and Conscientiousness. Openness to experience/Intellect reflects the engagement with aesthetic/sensory and abstract/intellectual information, as well as the degree of appreciation and toleration for the unfamiliar (Nicholson et al. 2002; Fleischhauer et al. 2010; Fayn et al. 2015). Extraversion relates to approach behavior of driving toward a goal that contains cues for reward, and tendency to experience positive emotions given by the actual attainment of that goal (Depue and Collins 1999; DeYoung 2015). Neuroticism relates to a person's emotional life and reflects the tendency to heightened emotional reactivity to negative emotions (Goldberg and Rosolack 1994; Rusting and Larsen 1997; Gray and Mcnaughton 2000). Agreeableness relates to interpersonal behavior and reflects the degree of avoidance of interpersonal conflicts (stability between individuals) (Graziano et al. 2007; Butrus and Witenberg 2013). Conscientiousness reflects the degree to which individuals perform tasks and organize their lives, exhibiting a tendency to show self-discipline, act dutifully, and aim for achievement (stability within individuals) (Ozer and Benet Martínez 2006; Roberts et al. 2009) (cf. for more details McCrae and Costa 2004; DeYoung and Gray 2009).

Since the FFM of personality is based on language descriptors of adjectives applied to human and human behaviour in English lexicon, rather than neurobiological features, many attempts have been made to explore the neural bases of these five factors. At first, each trait has been associated to its most crucial and characterizing psychological functions (e.g. Neuroticism and Extraversion to sensitivity to punishment and reward respectively, Agreeableness to social processes, Conscientiousness to top-down control of behaviour and Openness cognitive flexibility), and hypotheses have been developed about the associations between brain systems supporting those psychological functions, and the respective trait, paving the way for a biology of personality traits (c.f. DeYoung and Gray 2009). It has, therefore, been suggested that Neuroticism is associated (functionally or structurally) to affective regions that had been linked to respond to threat and punishment like amygdala, hippocampus, cingulate cortex and medial prefrontal cortex (Kumari 2004; Cremers et al. 2010; DeYoung et al. 2010; Tzschoppe et al. 2014; Madsen et al. 2015; Pang et al. 2016). Extraversion has been linked to regions responding to reward-related stimuli like nucleus accumbens, striatum, amygdala and orbitofrontal cortex (DeYoung et al. 2010b; Adelstein et al. 2011; Pang et al. 2016, c.f. Lei et al. 2015). Conscientiousness has been related to the lateral prefrontal cortex (Asahi et al. 2004; Passamonti et al. 2006; DeYoung et al. 2010; Kunisato et al. 2011), deputed to the planning, following complex rule and voluntarily control of

behavior. Similarly, Openness has also been associated to the functions of the lateral PFC (DeYoung et al. 2005; Kunisato et al. 2011), but in contrast to Conscientiousness, more because of its role in attention, working memory and cognitive flexibility. Finally, Agreeableness has been associated to regions involved in the processing of social information, such as temporo-parietal junction, superior temporal gyrus and posterior cingulate cortex (Hooker et al. 2008; DeYoung et al. 2010; Adelstein et al. 2011). However, the associations between brain systems underlying specific mental functions and personality traits might be more complex than such one-to-one mapping; instead, it is much more plausible that the mapping between traits and brain systems is rather many-to-many (c.f. Yarkoni 2015; Allen and DeYoung 2016). One example is provided by Neuroticism, which has not only been associated to affective regions, but also to regions exerting cognitive functions, e.g. dlPFC (Kunisato et al. 2011; Pang et al. 2016), or behavioural performances probing attention (MacLean and Arnell 2010), working memory (Studer-Luethi et al. 2012), verbal fluency (Sutin et al. 2011) and explicit memory (Pearman 2009; Denkova et al. 2012). It is therefore possible that these systems (affective and executive) both contribute in explaining variance in Neuroticism. The potential contribution of other regions rather than the ones originally suggested also holds for other traits. For example, increasing evidence points to a link between Openness and the functional organization and global efficiency of the default mode network (DeYoung 2014; Sampaio et al. 2014; Beaty et al. 2016). Similarly, even if not directly investigating the trait of Agreeableness, there is evidence (Gazzola et al. 2006; c.f. Iacoboni 2009) showing a possible association between one of its facet, empathy, with the mirror neuron system.

Furthermore, one of the major challenges of using functional studies for the association between personality traits and brain systems is the fact that the latter can only be based on specific implementations such as behavioural tests or paradigms used in experimental research. Moreover, there is a general consensus that mental functions arise from the coordinated activity within distributed networks rather than any individual brain region (Eickhoff and Grefkes 2011). Therefore, relating a personality trait to a particular function only because a brain region correlates with both is problematic. These considerations have prompted a network-centred perspective of brain organization (c.f. De Vico Fallani et al. 2014), highlighting the importance of functional integration for mental processes and their inter-individual differences. However, this approach, which requires a priori defined seeds, suffers from an important methodological limitation. That is, by choosing pre-defined nodes from a single task-based fMRI study, the findings might be biased toward that particular paradigm operationalization. Furthermore, task-based fMRI literature often suffers from low statistical power and low reproducibility, due to the small sample sizes typically used and considerable heterogeneity in the analysis pipeline (cf. Samartsidis et al. 2017). To solve the problem of a more objective definition of relevant nodes in a given functional network, quantitative meta-analyses of task-based neuroimaging studies aggregate the findings of

many individual task-activation studies into a core network representing those locations that are reliably recruited by engaging in a given kind of mental process (cf. Fox, Lancaster, Laird, & Eickhoff, 2014). The investigation of RSFC in meta-analytically defined networks representing specific social, affective, executive, or memory functions, therefore, provides a viable approach to capturing the complex intrinsic neural architecture underlying personality (Adelstein et al. 2011; Sampaio et al. 2014).

Given that network connectivity data are almost inevitably high-dimensional, consisting of many correlated features, univariate analyses of associations between connectivity measures and phenotypical traits such as personality may not represent an optimal strategy (Orrù et al. 2012). Moreover, univariate analyses will likely fail to elucidate associations that depend on the pattern of connectivity within a network rather than any specific individual connection. On the other hand, machine learning and multivariate pattern analysis (MVPA), suitable for analysing neuroimaging data (cf. Oktar & Oktar, 2015; Gael Varoquaux & Thirion, 2014), provide an approach that overcomes these limitations by searching for patterns in the connectivity matrix that allow the prediction of a continuous target variable (Doyle et al. 2015). In this article, the term "prediction" refers to the out-of-sample evaluation of a statistical model's ability to predict the personality score for previously unseen individuals based on their RSFC. The potential of such approaches to predict behavioural scores from resting-state connectivity data has already been demonstrated with respect to sustained attention (Rosenberg et al. 2016), autistic traits (Plitt et al. 2015) and impulsivity in economic decision-making (Li et al. 2013). Conversely, personality traits have been predicted from cyber records such as personal web sites (Marcus et al. 2006) or social networks (Golbeck et al. 2011; Bachrach et al. 2012) but not yet from neuroimaging data.

Bringing together the different aspects outlined above, the current study explored whether individual levels of five major personality traits can be predicted from RSFC profiles in a priori defined brain networks representing specific cognitive functions. The selection of the networks used a priori knowledge based on the associations reported in literature between psychological functions (and deputed networks) with personality. Accordingly, we chose functional networks associated to affective (emotion processing, reward and pain) functions given their main associations with both Extraversion and Neuroticism, social (empathy and face processing) functions in relation to Agreeableness, executive functions as linked to Conscientiousness and Openness (vigilant attention and working memory to represent respectively rigid control and flexibility) and memory (autobiographic and semantic) functions as many traits were also found to be associated with them. However, it is important to note that we refrained from having hypotheses about network – predicted traits associations, since we believe that multiple brain systems, among the selected ones, can contribute to explaining inter-individual variance in one trait (e.g. Openness being predicted from networks outside the executive domain). We additionally used a network with whole-

brain coverage consisting of 264 nodes (we here refer to it as **Connectome**; Power et al. 2011) to predict the five personality traits in order to test if personality can be better predicted by specific functional networks or a rather unspecific whole-brain network. Additionally, in light of previous findings of sexual dimorphism in the relationships between brain structure and personality traits (Nostro et al. 2016) as well as gender differences in RSFC (Allen et al. 2011; Filippi et al. 2013; Hjelmervik et al. 2014; Weis et al. 2017) and personality (Yang et al. 2015), these analyses were performed in a gender-mixed sample as well as separately in male and female subsamples.

2 Materials and methods

2.1 Participants

All data were obtained from the Human Connectome Project (HCP) WU-Minn Consortium as provided in the current "S1200" release (http://www.humanconnectome.org (Van Essen et al. 2013). The HCP was funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University. Our analyses of the HCP data were approved by the ethics committee of the Heinrich Heine University Düsseldorf.

The HCP sample is composed of monozygotic and dizygotic twins as well as not-twins, the latter including siblings of twins, just siblings, and only-children (including those that have an as-yet not scanned sibling but not twin). Given this structure of related and unrelated subjects, we paid particular attention to select a well-matched sample of males and females that was as large as possible while at the same time controlling for possible effects of heritability by creating a sample of only unrelated subjects. Evidently, we first selected all participants from the HCP sample for whom resting-state fMRI volumes and personality data were available. Out of this sample, we then selected groups of unrelated males and females (i.e. only one representative of a given family), matched for age, year of education and twin-status. This last match (twin or not twin) was preferred over the match for zygosity (not twin, dizygotic or monozygotic) as it enabled us to select a higher number of participants while not introducing dependencies in the sample. In fact, Kolmogorov-Smirnov test showed that zygosity does not lead to any significant difference in the five scores distribution, cf. **supplementary Table S1**. Importantly, we created a first main sample (**Sample 1**), where we aimed for the highest number of participants according to the inclusion criteria, but since a considerable number of individuals were left out from the first selection, we additionally created a "replication" sample, (**Sample 2**). **Sample 2** was thus created by removing the subjects belonging to the **Sample 1** from the main release (S1200) and re-applying the selection criteria on the remaining participants. The final selection procedure of **Sample 1** resulted in a total of 420 subjects: 205 males (119 non-twins, 91 twins subjects; aged 22-37 years, mean: 28.3 ± 3.5; years of education: 14.9 ± 1.8) and 205 females (117 non-twins, 93 twin subjects; aged 22-36 years, mean: 28.8 ± 3.5; years of education: 15.0 ± 1.8). **From the remaining subjects not selected for Sample 1, Sample 2** was obtained resulting in a sample of 302 subjects: 151 males (75 non-twins, 76 twins subjects; aged 22-36 years, mean: 28.2 ± 3.4; years of education: 14.8 ± 1.8) and 151 females (76 non-twins, 75 twin subjects; aged 22-35 years, mean: 28.9 ± 3.5; years of education: 15.0 ± 1.8). For an overview on the samples selection, see **Fig 1**.
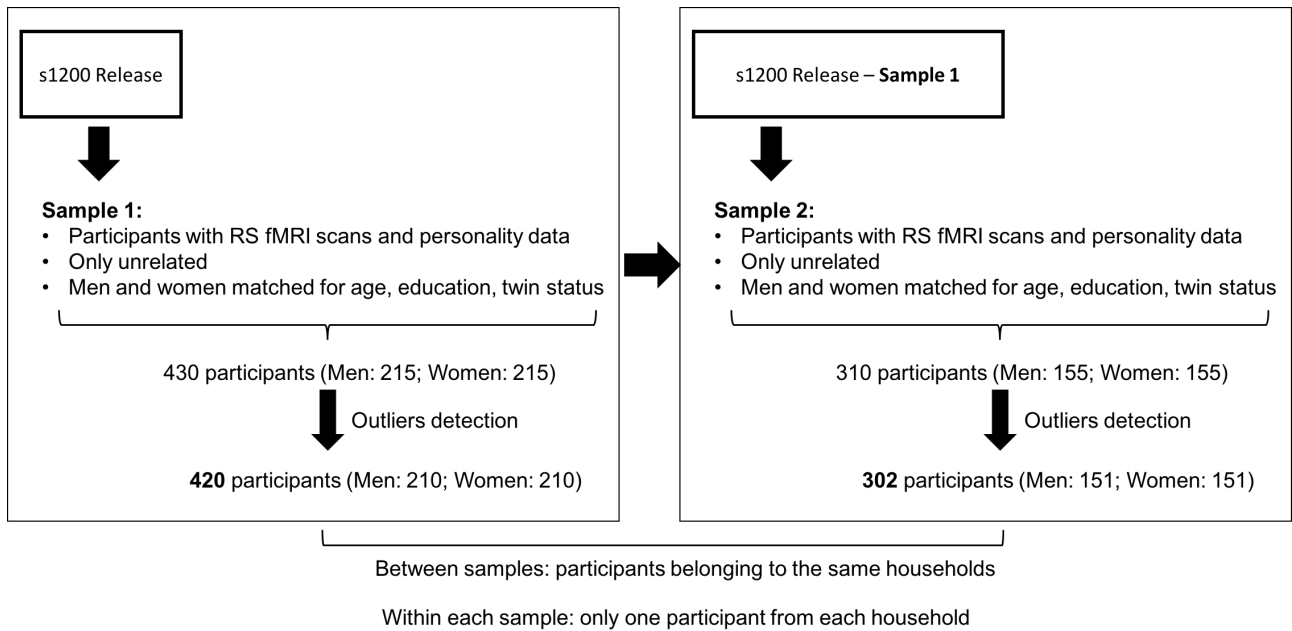
Fig 1: Samples selection overview: first Sample 1 (or "main" sample) was created aiming for the largest number of participants. Once 430 subjects were selected for this sample, the same procedure was applied on the remaing subjects of the HCP to generate Sample 2 (or "replication" sample). The two samples result in this was related to each other (as siblings of the subjects in Sample 1 are present in Sample 2), but, within each sample, there are no subjects related to each other.

2.2 Self-report data

Personality was assessed using the English-language version of the NEO Five Factor Inventory (NEO-FFI; McCrae and Costa 2004). The NEO-FFI consists of 60 items in the form of statements describing behaviours that are characteristic for a given trait, 12 for each of the five factors (Openness, Conscientiousness, Extraversion, Agreeableness and Neuroticism). Each factor is assessed by aggregating individual responses given on five-point Likert-type ratings scales, yielding sum scores between 0 and 60 for each factor. Data were analysed using SPSS 20 (IBM Corp. Released 2011); scores of males and females were compared via t-tests ($p < 0.05$, Bonferroni-corrected for multiple comparisons) for each personality trait. In case of significant group differences, we estimated effect sizes by using Cohen's d measure (Cohen 1988). Furthermore, correlations among factors were calculated and tested for significance (Bonferroni-corrected) separately for males and females (for details, see **supplementary material**). Importantly, as reported on the HCP listserv (https://www.mail-archive.com/hcp-users@humanconnectome.org/msg05266.html), the Agreeableness factor score in the HCP database was erroneously calculated due to item 59 not reversed. We addressed this issue by reversing it and using the correct score of Agreeableness.

2.3 Meta-analytically derived networks

2.3.1 Selection of networks

We selected nine meta-analytic networks representing regions consistently activated by various social, affective, executive and memory functions. Specifically, we used two networks related to social cognition: empathy (**Emp;** Bzdok et al., 2012) and static face perception (**Face;** Grosbras, Beaton, & Eickhoff, 2012); three networks related to affective processing: reward (**Rew;** Liu, Hairston, Schrier, & Fan, 2011), physiological stress/pain (**Pain;** Kogler et al., 2015) and perception of emotional scenes and faces (**Emo;** Sabatinelli et al., 2011); two networks related to executive functions: working memory (**WM;** Rottschy et al., 2012) and vigilant attention (**VA;** Langner & Eickhoff, 2013); two networks related to long-term memory: autobiographic memory (**AM;** Spreng, Mar, & Kim, 2008) and semantic processing (**SM;** Binder, Desai, Graves, & Conant, 2009).

2.3.2 Selection of coordinates

From each meta-analysis, we selected the reported coordinates of the networks to include in our analyses and modelled a 6-mm sphere around each coordinate. This ensured that all nodes were represented by region of interest of equal size (ROIs) within and across networks. Within each single network, we only selected peaks that either represented different anatomical regions, preventing multiple representations of a single region, or were at least 15 mm apart from each other (according to the SPM anatomy toolbox 2.1; (Eickhoff et al. 2005, 2007)). In cases of multiple peaks within an anatomical region that were closer to each other, we included the peak showing the highest Z-score. Please note, these criteria were only applied for multiple regions within a single network, while we did not exclude any regions that were found also in another network. That is, even if different networks featured peaks at the same location, these presumably shared nodes were retained. Given that little is yet known about the effect of the networks' sizes on the outcome predictability, we also had to consider the size of the networks (i.e. number of nodes) to make sure that possible differences in their predictive power were not due to the number of nodes included. As a result, the size of the networks ranged between 16 (**VA**) and 24 (**Emo**) nodes. Further details on the meta-analytic networks can be found in **Table 1**, **supplementary Table S3** and **supplement Fig S1**.

Table 1: Description of the meta-analytic derived networks

Domain	Meta-analytic Network	Abbreviation	Author, Year	Reference of the network in the original paper	Number of included Nodes	Network description
Social	Empathy	**Emp**	Bzdok, 2012	Table n.1 (ALE meta-analysis of empathy)	22	Regions consistently activated during tasks referring to conscious and isomorphic experience of somebody else's affective state
Social	Static Face Perception	**Face**	Grosbras, 2012	Table n. 7 (Static face perception)	19	Convergence across tasks consisting in viewing photographs of faces or viewing objects/ scrambled images
Affective	Reward	**Rew**	Liu, 2011	Table n. 1	23	Convergence across reward valence and decision stages contrasts
Affective	Physiological Stress	**Pain**	Kogler, 2015	Table n.1 (Activation physiological)	18	Regions consistently activated during tasks referring to unpleasant sensoric, emotional and subjective experience that is associated with potential damage of body tissue and bodily threat
Affective	Perception of emotional scenes and faces	**Emo**	Sabatinelli, 2012	Table n.2 (emotional face>neutral face) & Table n.3 (emotional scenes>neutral scenes)	24	Regions consistently activated during tasks referring to discrimination of emotional faces> neutral faces contrast combined with emotional scenes> neutral scenes contrast

Executive	Working Memory	**WM**	Rottschy, 2012	Table n. 2	22	Regions consistently activated during all WM contrasts/ experiments (mainly n-back, Stenberg, DMTS, delayed simple matching)
Executive	Vigilant Attention	**VA**	Langner, 2012	Table n.1	16	Regions consistently activated during tasks posing only minimal cognitive demands on the selectivity and executive aspects of attention for more than 10s
Memory	Autobiographic Memory	**AM**	Spreng, 2008	Table n. 6	23	Convergence across tasks referring to autobiographical recall: episodic recollection of personal events from one's own life
Memory	Semantic Memory	**SM**	Binder, 2009	On request to the author	23	Regions consistently activated during all SM contrasts/ experiments (mainly words vs. pseudowords, semantic vs. phonological task, high vs. low meaningfulness)
Whole-brain	Connectome	**Connectome**	Power, 2011	Supplement material	264	Meta-analytic ROIs and FC-mapping ROI merged to form a maximally-spanning collection of ROIs. Meta-analytic ROIs were given preference, and non-overlapping fc-mapping ROI were then added

2.4 Connectome analysis

In addition, we employed a brain-wide network of 264 functional areas from Power and colleagues (**Connectome**; Power et al. 2011) to compare the predictive power of RSFC from the whole-brain and from meta-analytic networks. For the coordinates of this **Connectome**, please refer to the supplementary Table S2 of Power et al.

2.5 Resting-state fMRI data: Acquisition, preprocessing and functional connectivity analyses

As part of the HCP protocol (Glasser et al. 2013), images were acquired on a Siemens Skyra 3T Human Connectome scanner (http://www.humanconnectome.org/about/project/MR-hardware.html) using a 32-channel head coil. Resting-state (RS)-BOLD data (voxel size= 2 x 2 x 2 mm³, FoV= 208 x 180 mm², matrix = 104 x 90, 72 slices in a single slab, TR = 720 ms; TE= 33.1 ms, flip angle = 52°) were collected using a novel multi-band echo planar imaging pulse sequence that allows for the simultaneous acquisition of multiple slices (Xu et al. 2013). RS-fMRI data were then cleaned of structured noise through the Multivariate Exploratory Linear Optimized Decomposition into Independent Components (MELODIC) part of FSL toolbox (www.fmrib.ox.ac.uk/fsl). This process pairs independent component analysis with a more complex automated component classifier referred to as FIX (FMRIB's ICA-based X-noisifier) to automatically remove artefactual components (Salimi-Khorshidi et al. 2014).

The FIX-denoised RS-fMRI data were further preprocessed using SPM12 (Statistical Parametric Mapping, Wellcome Department of Imaging Neuroscience, London, UK, http://www.fil.ion.ucl.ac.uk/spm/), running under Matlab R2016a (Mathworks, Natick, MA). For each participant, the first four EPI images were discarded prior to further analyses. Then EPI images were corrected for head movement by affine registration using a two-pass procedure: in the first step, images were aligned to the first image, and in the second step to the mean of all volumes. Next, the mean EPI image was spatially normalized to the non-linear MNI152 template (Holmes et al. 1998) by using the "unified segmentation" approach in order to account for inter-individual differences in brain morphology (Ashburner and Friston 2005). Finally, images were smoothed with an isotropic Gaussian kernel (full-width at half-maximum = 5 mm). The activity time series of each voxel was further cleaned by excluding variance that could be explained by mean white-matter and cerebrospinal-fluid signal (Satterthwaite et al. 2013). Data were then band-pass filtered with cut-off frequencies of 0.01 and 0.08 Hz.

In order to identify participants with aberrant RSFC patterns, we computed each subject's entire connectome sampled on a 1-cm grid. We then computed the pairwise Euclidean distance between the subjects and identified the nearest neighbour for each subject. We excluded the subjects whose distance to their nearest neighbour was in the highest 2.5% and

at least 3 SD away from the average distance. This procedure was done separately for men and women (Sample 1: 5 males, 5 females; Sample 2: 4 males, 4 females). No subjects were excluded due to outlier motion parameters (DVARS and FD both displaying zero-centered values) (Salimi-Khorshidi et al. 2014; Varikuti et al. 2016; Ciric et al. 2017). For RSFC analyses, the subject-specific time series for each node of each network were computed as the first eigenvariate of the activity time courses of all grey matter voxels within 6 mm of the respective peak coordinate. We then computed pairwise Pearson correlations between the eigenvariates of all nodes in each network, which then were transformed using the Fischer's Z scores and adjusted (via linear regression) for the effects of age and movement.

2.6 RSFC-based prediction of personality traits by Relevance Vector Machine learning

We examined if the RSFC patterns within each network predicted personality scores by means of statistical learning via the Relevance Vector Machine (RVM; Tipping, 2001) as implemented in the SparseBayes package (http://www.miketipping.com/index.htm). The RVM is a machine learning technique that can learn to predict a continuous target value given explanatory variables (also called features). In our case the features were the RSFC values between all nodes of a meta-analytic network, while the score of a specific personality factor scale was the target value.

Briefly, RVM is a multivariate approach that was developed from the Support Vector Machine (SVM) in order to induce sparseness in the model's parameters. The RVM, in contrast to SVM, implements a fully probabilistic Bayesian framework: for each possible value of the input vector (e.g. set of FC values), the RVM algorithm provides a probability distribution of the predicted target value (e.g. FFM personality score), unlike a point estimate obtained by the SVM.

$$\hat{y}(x, w) = w_0(0; \sigma_0) + \sum_{i=1}^{n} w_i (0; \sigma_i) K_\sigma(x_i, x),$$

In the RVM formulation above, the kernel K is a multivariate zero-centered Gaussian with standard deviation σ (estimated by the algorithm) and every parameter w_i, assigned to each subject x_i in the training set, is assumed to follow a Gaussian with mean zero and standard deviation σ_i. The standard deviations σ_i that describe the probability distribution of the parameters w_i are iteratively estimated from the training data in order to maximize the likelihood of the model. Sparseness is achieved by discharging parameters w_i converged to zero. Once σ_0 and σ_i have been estimated, the trained model can be used to predict the target value (e.g., FFM personality score) from a previously unseen input vector (RSFC data from participants that were not part of the training data) by computing the predictive distribution (for a more detailed description, see Tipping, 2001).

In our study, we implemented the RVM algorithm with a 10-folds cross-validation. That is, the sample was randomly split into 10 equally sized groups of which 9 were used for training while one was held back and used for assessing the performance of the prediction in previously unseen data. Holding out each of the 10 groups in turn then allowed computing the prediction performance across the entire dataset. Importantly, this procedure was repeated 250 times using random initial splits of the data to obtain robust estimates of the RVM performance for predicting a given NEO-FFI score from a particular network's RSFC pattern. For each subject, the predicted values resulting from each cross-validation (i.e. one replication) were averaged over the 250 replications and ultimately correlated with the real score. As we performed 250 replications of a 10-fold cross-validation, in total 2500 models were computed to predict each trait. We thus quantified the contribution of each connection by the fraction of these 2500 models in which the weight for the respective connection was non-zero. The connections that had a non-zero weight in at least 80% of all models were identified as the connections that were most robustly part of the predictive model. The brain networks were visualized with the BrainNet Viewer (http://www.nitrc.org/projects/bnv/) (Xia et al. 2013).

For both the "main" (**Sample 1**) and "replication" (**Sample 2**) samples, predictions were first carried out for all subjects with males and females combined (All$_{Sample1}$: n = 410 All$_{Sample2}$: n = 302), and then separately for the male (Men$_{Sample1}$: n = 210; Men$_{Sample2}$: n = 151) and female group (Women$_{Sample1}$: n = 210; Women$_{Sample2}$: n = 151) in order to assess gender differences in predictability. Predictive power was assessed by computing Pearson correlations between real and predicted NEO-FFI scores and mean absolute error (MAE). Importantly, results were only regarded as significant when they were significant at a threshold of $p < 0.05$ in **both** samples (Sample 1 and Sample 2). The p value was computed via permutation testing between real and predicted values with 10.000 runs. For each run, we shuffled the predicted scores across subjects in either the entire sample (for "All") or in the gender-groups (for "Men" and "Women") without replacement. From here, the definition of the p value as the fraction of runs when the correlation between real and the shuffled predicted score was higher than the one obtained between the real and the original predicted value.

For all significant results in either "All", "Men" or "Women", we further tested for significant differences in prediction performance (i.e. correlation between real and predicted value) between males and females in the main sample. Pearson correlation coefficients (r) were transformed into Fisher's Z and the difference between Z_{Men} and Z_{Women} calculated and then 95% confidence intervals (CI) were computed based on these difference scores. The difference in correlation coefficients between males and females were regarded as significant if the 95% confidence interval did not contain zero (Lane 2013).

3. Results

3.1 NEO-FFI scores

Subjects scored in the same range as reported by McCrae and Costa (McCrae and Costa 2004) in both the samples.

Correlations between factors were calculated separately for males and females and in the entire sample (see **supplementary Table S2** for more detailed information). Most of them were significant at $p < 0.05$ (Bonferroni-corrected) in both males and females and the entire sample. Openness, however, was found to be independent of most of the other factors, except for Agreeableness (in **Sample 1** for All, Men and Women), and Conscientiousness (in All for both **Sample 1** and **Sample 2**). Furthermore, Neuroticism was the only factor correlating negatively with almost all the others (except for Openness in Men of **Sample 1** and in All, Men and Women of **Sample 2**).

Comparison of the scores for the five personality traits between Men and Women revealed a significant difference for Agreeableness in both samples (**Sample 1**: $t_{407} = -4.95$; $p < 0.05$, d = -0.49; **Sample 2**: $t_{299} = -2.2$; $p < 0.05$, d = -0.27), with females scoring higher than males. For Neuroticism, Women significantly scored higher than Men in **Sample 1** ($t_{407} = -2.8$; $p < 0.05$, d = -0.28), while in **Sample 2** this difference only showed a trend ($t_{299} = -1.93$; $p = 0.055$, d = -0.2). For Openness (**Sample 1**: $t_{407} = 0.1$; $p = 0.9$; **Sample 2**: $t_{299} = 1.64$; $p = 0.1$) and Extraversion (**Sample 1**: $t_{407} = 1.1$; $p = 0.3$; **Sample 2**: $t_{299} = -0.68$; $p = 0.5$) no significant gender differences were found. For Conscientiousness, Women significantly scored higher than Men in **Sample 2** ($t_{299} = -2.11$; $p < 0.05$, d = -0.245), while in **Sample 1** Women scored higher than Men, but not significantly ($t_{407} = -0.41$; $p = 0.15$).

3.2 RVM: Predicting personality traits based on RSFC

Results are only be reported if they were significant both in the main (**Sample 1**) and in the replication sample (**Sample 2**).

3.2.1 Predictions in the entire sample (balanced males & females)

In the entire sample, the RSFC pattern of four networks significantly predicted personality factors: **Pain** and **VA** predicted Openness, **AM** predicted Agreeableness and **Connectome** predicted Neuroticism (see **Table 2**, **Fig 2** for an overview of the results and **Fig 3** for the correlation plots).
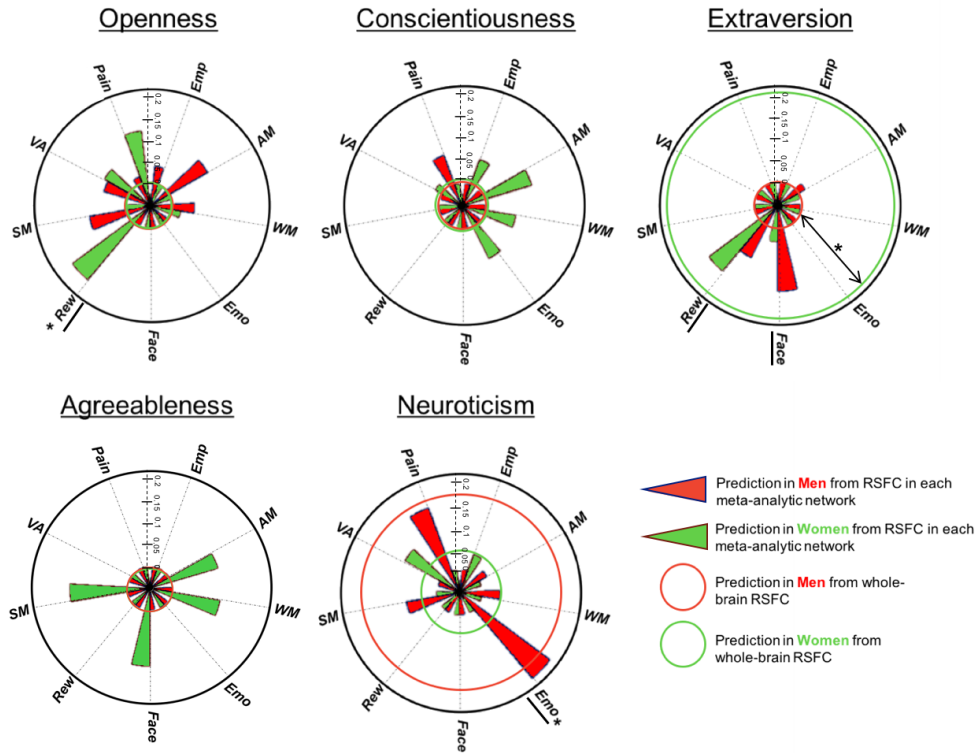
Fig 2: Emp: empathy; **AM**: Autobiographic memory; **WM**: working memory; **Emo**: emotional processing; **Face**: face processing; **Rew**: reward; **SM**: semantic memory; **VA**: vigilant attention; **Pain**: pain processing.

Summary of the networks for which FC patterns significantly predicted the five personality traits. For each network-trait combination in either Men or Women, here it is reported the conjunction between the correlation coefficients (i.e. minimum r value). Only predictions with r > 0.1 are displayed. While the nine meta-analytic networks are represented as slices (triangules) of the five personality circles, the connectome is represented as well as a circle. Triangules and circles are scaled based on the r values of the predicting networks (r values reported in the axis). Meta-analytic networks are underlined if a significant prediction is detected in either Men or Women. Asterisks mark significant gender differences in **Sample 1**.
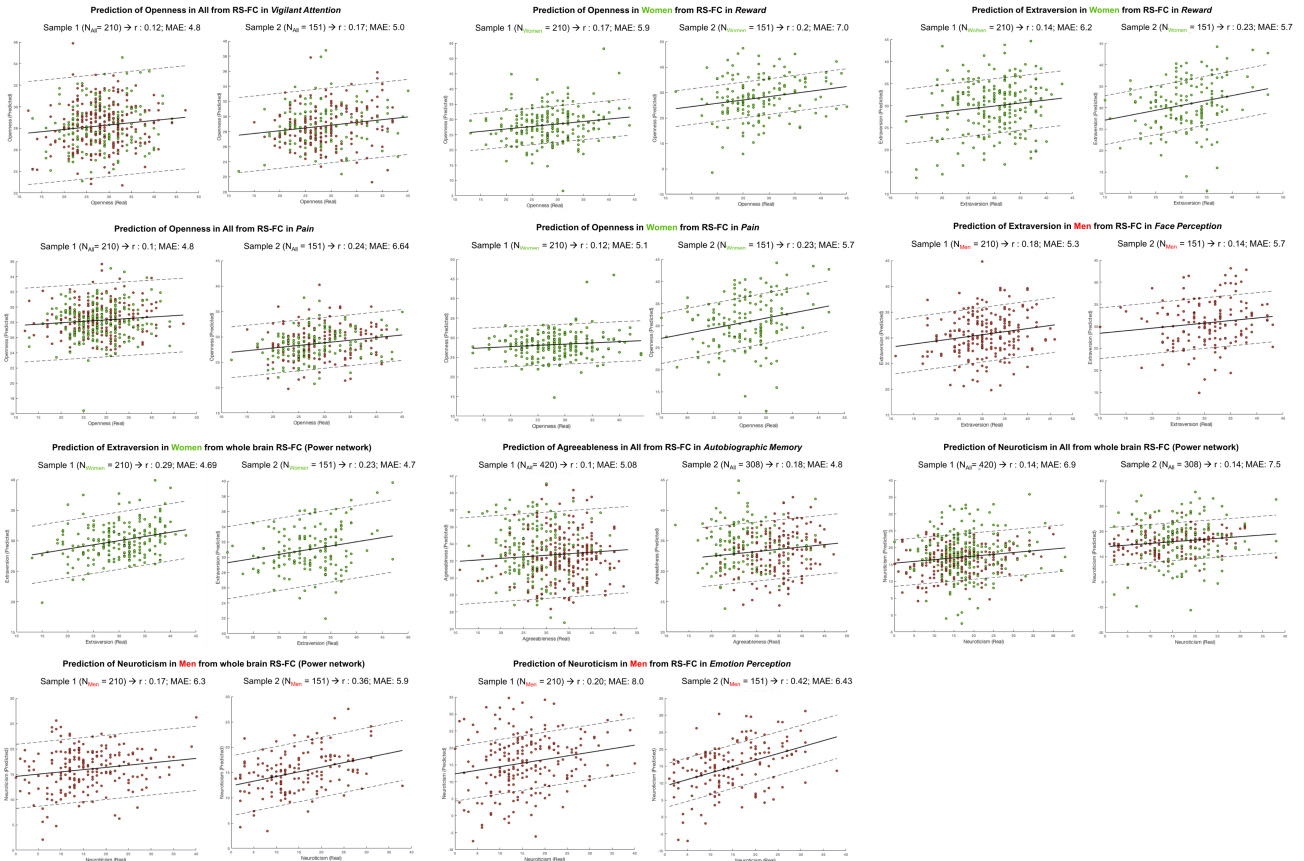
Fig 3: Scatter plots of the predictions of personality scores significant at p < 0.05 in both samples. Continuous regression lines, dashed lines, representing the standard deviation, and mean absolute errors (MAE) are displayed.

Table 2: Results of the Relevance Vector Machine

Predicted Trait	Predicting Network	Group	r (Sample1)	p-value (Sample1)	r (Sample2)	p-value (Sample2)
O	VA	All	0.12	0.006	0.17	0.001
O	Pain	All	0.1	0.018	0.2	0.0
O	Rew	Women	0.17	0.006	0.2	0.006
O	Pain	Women	0.12	0.048	0.29	0.0
E	Face	Men	0.18	0.005	0.14	0.04
E	Rew	Women	0.14	0.02	0.23	0.002
E	Connectome	Women	0.29	0.0	0.23	0.002
A	AM	All	0.1	0.018	0.18	0.001
N	Connectome	All	0.14	0.018	0.14	0.04
N	Connectome	Men	0.17	0.0	0.38	0.0
N	Emo	Men	0.2	0.002	0.42	0.0

Predicted Trait: O: Openness; E: Extraversion; A: Agreeableness; N: Neuroticism.

Predicting Network: **VA**: vigilant attention; **Pain**: pain processing; **Rew**: reward; **AM**: autobiographic memory; **Face**: face perception; **Connectome**: whole-brain network; **Emo**: emotional processing.

Correlation coefficients between real and predicted values which resulted significant at $p < 0.05$ in **both** samples in either across the entire sample (All), or in gender groups (Men or Women).

3.2.2 Predictions of personality traits in the gender-split groups

In the gender-split groups, we also found a significant prediction of Openness scores based on FC patterns within the **Pain** network in Women as well as prediction of Neuroticism based on the **Connectome** FC in Men. In contrast, the **VA** and **AM**-related networks did not significantly predict Openness and Agreeableness in either subgroup. However, in the gender-specific groups additional significant predictions were observed: in males, Extraversion was predicted by the RSFC patterns of **Face** and Neuroticism by **Emo** networks (**Table 2**, **Fig 2-3**). In females, Openness was predicted by **Rew** network. Furthermore, in females, Extraversion was predicted by **Rew** network and the **Connectome** (**Table 2**, **Fig 2-3**).

3.3 Gender differences in personality predictability

For all the predictions that were significant in at least one group (All/Males/Females), we tested if prediction performance was significantly different between the male and female subgroups. Significantly better predictability in Men than Women was found for Neuroticism predicted from **Emo** network (**Table 3**, **supplementary Fig S2**). In Women compared with Men, Openness was significantly better predicted from **Rew** network and Extraversion from the entire **Connectome** (**Table 3**, **supplementary Fig S2**).

Notably, not all associations that were only found predictive in one subgroup showed significant differences in predictability between males and females. In particular, no gender differences were found in predicting Openness from **Pain**, and **VA** networks, Neuroticism from **Connectome**, Agreeableness from **AM**, and Extraversion from **Face** and **Rew** networks (**Table 3, supplementary Fig S2**).

Table 3: Gender differences in personality predictability

Predicted Trait	Predicting Network	Group	r (Sample1)	$Z_{Men} - Z_{Women}$ (Cohen's q)	CI (Lower limit / Upper limit)
O	VA	Men	0.06	0.013	-0.176 / 0.205
		Women	0.07		
O	Pain	Men	0.08	0.039	-0.153 / 0.231
		Women	0.12		
O	Rew	Men	-0.06	0.236 *	0.044 / 0.428
		Women	0.17		
O	Pain	Men	0.08	0.039	-0.153 / 0.231
		Women	0.12		
E	Face	Men	0.18	0.054	-0.138 / 0.246
		Women	0.12		
E	Rew	Men	0.08	0.055	-0.137 / 0.247
		Women	0.14		
E	Connectome	Men	-0.03	0.323 *	0.131 / 0.515
		Women	0.29		
A	AM	Men	0.10	0.190	-0.002 / 0.382
		Women	-0.09		
N	Connectome	Men	0.17	0.119	-0.073 / 0.311
		Women	0.06		
N	Emo	Men	0.2	0.276 *	0.084 / 0.468
		Women	-0.07		

Comparison of the correlation coefficients between males and females and effect size of significant gender differences. Confidence intervals (CI) are computed on the Z-transformed difference between correlations in Men and Women for each prediction. Note * marks significant gender difference at 95% of confidence.

4.   Discussion

Here we report associations between major dimensions of personality and RSFC in functional brain networks. In particular, individual scores of various personality traits of the Five-Factor Model (McCrae and Costa 2004) could be predicted from patterns of RSFC in specific meta-analytically defined networks as well as from the whole-brain FC pattern. In assessing the generalizability of our findings, we focused on the predictions that replicated in two different samples within the HCP dataset.

These results capitalize on the as-yet largely untapped potential (though cf. Schilbach et al., 2016; Varikuti et al., 2016) of neuroimaging meta-analyses to provide robust, functionally specific ROIs to investigate individual task-free data (Lee et al. 2012). These can help to constrain the otherwise vast feature space for statistical learning on resting-state data in a functionally meaningful and anatomically specific manner (Wang et al. 2010). As we demonstrate here, combining meta-analytic network definitions with statistical learning approaches allows, at a moderate level, not only predicting complex individual characteristics such as personality traits, but also the characterization of functional brain networks by their capability to do so. Nonetheless, our results of prediction of personality based on whole-brain FC pattern highlight that for some traits it might be crucial to consider the global connectivity as well.

In the overall (gender-mixed) sample, RSFC within networks representing affective and executive brain systems predicted Openness, RSFC within mnemonic network predicted Agreeableness, while RSFC from the whole brain predicted Neuroticism. In the gender-split samples, however, the prediction of Openness from the executive network **VA** and of Agreeableness from the mnemonic network **AM** were not replicated in any of the two subgroups, an effect likely related to the moderate effect present in the overall sample not specifically driven by a particular sex. In contrast, the prediction from the affective network **Pain** was also predicted in the female-only subsample, indicating that more information on the respective phenotypes can be gained from RSFC data in one gender. The gender-specific analyses revealed further constellations in which personality traits could be predicted from particular networks (see **Fig 2**). In fact, none of the network–trait combination was predictive in both female and male subsamples, but several functional networks were found to differentially predict personality traits in females versus males. Additionally, **Connectome** successfully predicted Extraversion (in Women) and Neuroticism (in the entire sample, but then also in Men only). This underlines the notion that gender is a fundamental factor with regard to brain–personality relationships.

4.1 Methodological considerations and limitations

In our analysis, we combined a priori selection of networks of interest, built upon the existing literature (cf. Kennis et al. 2013, Hu et al. 2011, DeYoung 2010), together with a data-driven approach for learning of the predictive models. The benefits of this approach were two-folds: on the one hand, with the a priori selection of networks, we could narrow down the networks of interest, which allowed us for a better functional interpretation of the results as the nodes represent brain regions robustly associated with the respective mental functions; on the other hand, the data-driven predictive models allowed for an explanatory analysis investigating which networks were informative in predicting a single trait, assuming therefore that many biological systems could contribute in explaining its inter-individual variance (Yarkoni 2015). Given that if only meta-analytically defined functional networks were employed, less consistently linked yet potentially critical regions might have been left out, we included also a purely explorative analysis employing the whole-brain FC.

In addition, as noted above, using a sparsity inducing method (RVM) which yielded compact regional modes has the advantage of providing regionally specific prediction models. As outlined above, our procedure provided a biologically informed feature reduction, as only the most relevant connections were taken in account in the prediction models. This has the advantage of reducing the complexity of the models avoiding overfitting (Hastie et al. 2009). With respect to the prediction model, we here employed Relevance Vector Machine (RVM), which in contrast to support vector regression or ridge regression, yields considerably sparser solutions (Tipping 2001). This allowed for identifying the most used connections and nodes (**Fig 4**) that mainly drove the prediction and hence enabled a more specific interpretation of its neurobiological underpinnings. In this context, it is important to note that for any given model the entire set of connections with non-zero coefficients provides information about the personality trait (Orrù et al. 2012). For interpretation, however, we focused on the most consistently utilized connections (over 250 replications) as key components of the given prediction.

In accordance with recent recommendations, the current study used 10-folds cross-validation, which has been showed to be less susceptible to overly optimistic estimates as compared with a leave-one-out approach (LOO-CV) (Varoquaux et al. 2016). Moreover, we repeated the cross-validation procedure 250 times, averaging the prediction performance over all replications to obtain robust and generalizable estimates of the capability of different brain networks to predict personality scores in new individuals.

Our approach, by building upon these methodological considerations, yielded insights into the relationships between brain, behaviour and personality. However, there are some limitations which are worth consideration in the future studies. First, gender-stratified sub-analyses may reduce statistical power because of the smaller sample sizes. Further studies

with a larger sample size, designed to separately analyze men and women are required, especially monitoring their hormonal levels (Arélin et al. 2015; Weis et al. 2017). Second, even though meta-analytic networks are among the most reliable ways to infer a mental function given a set of brain regions, we acknowledge that some regions of different functional networks can overlap. As a matter of fact, the employment of meta-analytically derived networks does not necessarily ensure a stringent and univocal relationship between the mental function supported by a particular network and a personality trait. Nonetheless, this approach can at least provide some confidence for the implication that a specific trait is related to a particular mental function in terms of the network that subserves them. A third consideration relates to the measurement of personality, i.e. the use of self-reported questionnaires. Self-reported questionnaire might have indeed contributed in increasing the noise in the data, as perception and report of own personality traits can be affected by many factors, e.g. men usually scoring low on Neuroticism as socialization effect (Viken et al. 1994).

4.2 Predicting Openness to experience

Our results indicated that self-reported Openness to experience can be linked to RSFC patterns in the networks subserving reward (**Rew**) and pain (**Pain**) processing in Women, while in the overall sample Openness was significantly predicted by RSFC in the vigilant attention (**VA**) network and, again, from **Pain**. Openness to experience has been linked to "need for cognition," that is, an individual's tendency to engage in effortful cognitive processing (Fleischhauer et al. 2010): high levels of Openness were found to positively affect work outcomes for highly complex jobs while increasing dissatisfaction when jobs become mechanical and unchallenging (Mohan and Mulla 2013). Such monotonous and intellectually unchallenging tasks were exactly the tasks investigated in the **VA** meta-analysis of Langner and Eickhoff (2013), which revealed the brain network involved in dealing with sustained attentional demands in boring situations. Thus, the predictability of Openness from FC in the **VA** network may reflect a neural substrate of the challenge experienced by individuals scoring high on Openness when faced with repetitive tasks and standardized routines. High-Openness participants might therefore need to recruit this network differently than their low-Openness individuals to keep focused on a tedious, repetitive task over time. Indeed, connections used throughout all prediction models from the **VA** network of Openness in both samples are between pre-supplementary motor cortex and medial prefrontal cortex (both involved in task-set re-energizing and outcome monitoring), between left inferior occipital gyrus (IOG) and right temporo-parietal junction (crucial for re-orienting the signalling), and left IOG and inferior frontal junction (known for its contribution in the input/output transformation) (see **Fig 4** for the most informative connections and Langner and Eickhoff 2013 for more details on the regions functions).

Behaviours associated with the trait of Openness, such as cognitive exploration, have been attributed to high dopamine (DA) functioning (DeYoung et al. 2005). This, indeed, led to the inclusion of Openness in the meta-trait "β" (or plasticity, c.f. DeYoung 2010), a higher order factor representing the shared variance between Openness and Extraversion, which are suggested to be both modulated by the dopaminergic system. DA is the main neurotransmitter modulating the reward network (cf. Berridge and Robinson 1998), and, in line with this, RSFC within the **Rew** network, could predict both Openness and Extraversion (in Women and in Men respectively), possibly via affecting the reactivity of the dopaminergic system. Interestingly, in predicting Openness, the weights of the nodes (i.e. number of incident edges) most used across the predictive models showed a stronger involvement of the dlPFC, corroborating previous findings that showed an association between Openness and the dopaminergic mesocortical branch, which projects directly onto the dlPFC (DeYoung 2013; Passamonti et al. 2015). On the other hand, regions like amygdala, nucleus accumbens (NAc) and orbitofrontal cortex (OFC), which constitute the other main dopaminergic branch, the mesolimbic pathway, were significantly less recruited. We would thus suggest that DA neurons populating the mesocortical branch, by encoding specifically the saliency of the stimulus (i.e. reward value of information, cf. Bromberg-Martin et al. 2010), can be potentially more informative for high-Open individuals, characterized by the automatic tendency to perceive salient information in everyday experience (DeYoung 2013). Interestingly, we found that Openness could be predicted by FC of the **Rew** network significantly better in Women, compared to Men (r = 0.17 in Women and r = -0.06 in Men of **Sample 1**). This might be explained by the fact that **Rew** functioning is highly influenced by the ovarian hormones estrogen and progesterone during the menstrual cycle (Dreher et al. 2007). In addition, estrogens have been related to dlPFC functioning, going along with cognitive decline which follows the drop of estrogens in menopause (Shanmugan and Epperson 2014). Despite the lack of studies exploring a direct relationship between females' hormonal cycling and the trait of Openness, there is evidence for its indirect modulation by estrogen. That is, the catechol-O-methyltransferase gene, which is associated with the trait of Openness (Konishi et al. 2014), is influenced by estrogen (Harrison and Tunbridge 2008). We thus suggest that the influence of ovarian hormones on RSFC in the **Rew** network as well as on perceived Openness induces joint intra-individual variation (i.e. shared variance), which in turn increases the strength of the neural and phenotypical association across women. This should then result in the observed higher predictability of Openness in female participants.

Across the entire sample, but then also in the female sub-group only, Openness could additionally be predicted in both samples based on FC within the pain network (**Pain**). Relationships between pain and Openness have been demonstrated in terms of a higher threshold for pain tolerance (Yadollahi et al. 2014) and as protective factor in migraine occurrence (Magyar et al. 2017) in individuals reporting higher levels of Openness.

However, very little is known about the association between this trait and the neural correlates of pain. Indirect evidence, however, comes from research in avoidance learning, which suggests that the successful avoiding of an aversive stimulus is experienced as an "intrinsic" reward (Kim et al. 2006). Endogenous opioid peptides, which are highly dense in the pain network (Baumgartner et al. 2006), were indeed found to modulate the dopaminergic system in response to aversive stimuli, resulting in the enhancement of a pleasure feeling boosted by DA (Sprouse-Blum et al. 2010). We thus suggest that high- and low-Open individuals differ in their ability to detect possible aversive stimuli (via diverse reactivity of the **Pain** network) and, by avoiding them, differently experience "intrinsic" reward.

In summary, the predictions from the **Rew**, **VA** and **Pain** networks of Openness might, therefore, jointly point to the importance of saliency processing of stimuli, which can be rewarding (**Rew**), monotonous (**VA**) or aversive (**Pain**), turning high Open-individuals as highly receptive and permeable to relevant information. Ultimately, connections between regions specially targeted by ovarian hormones (e.g, dlPFC), might underlie the significant gender difference in the predictability of Openness from FC in **Rew** network (**Fig 4**).
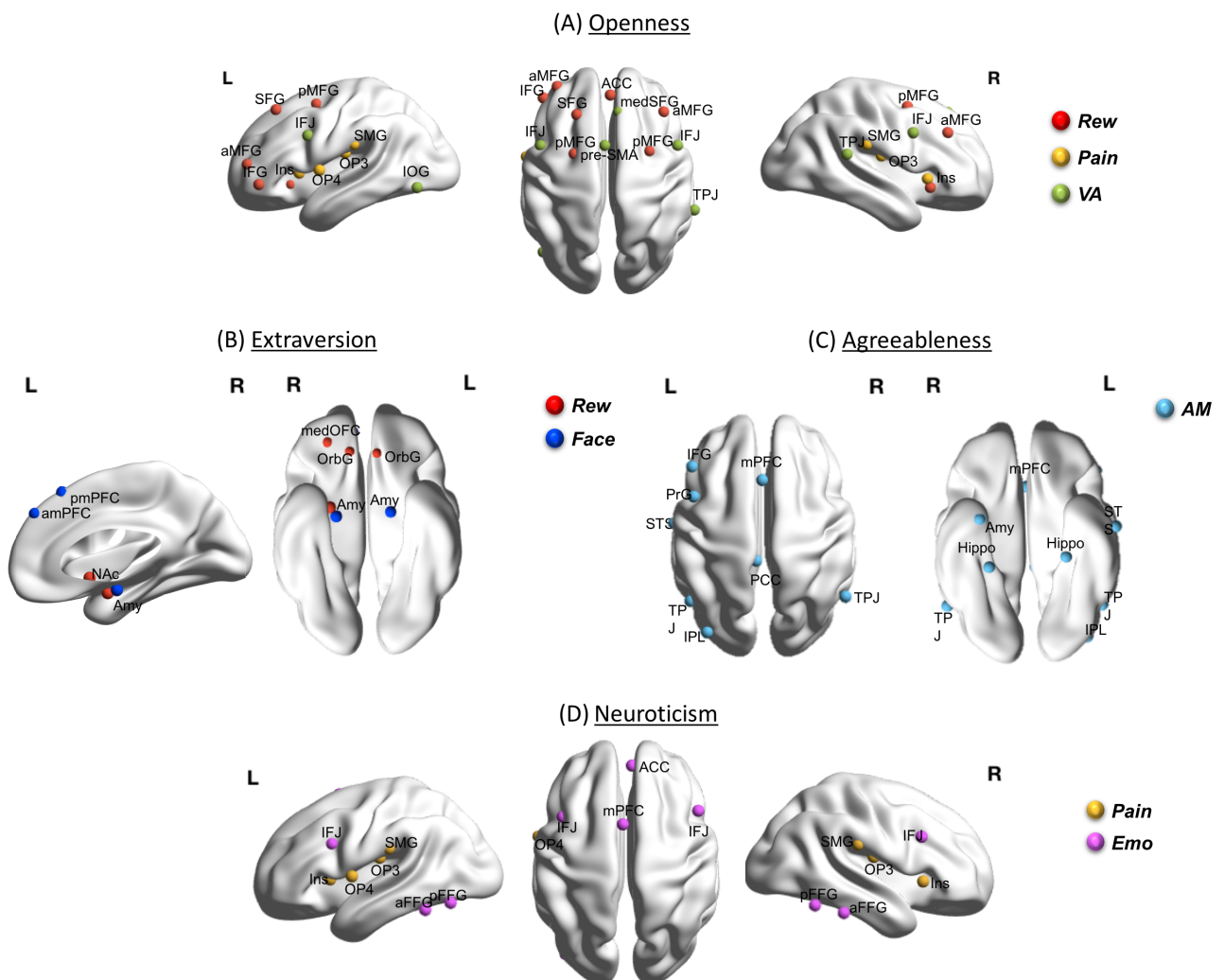

Fig 4: Summary of the most used nodes (i.e. above 80% of the models) between regions

from (A) the reward (**Rew**), vigilant attention (**VA**), and pain processing (**Pain**) networks in the prediction of Openness, (B) the Rew and face processing (**Face**) networks in the prediction of Extraversion. Summary of the most used connections between regions from (C) the autobiographic memory (**AM**) network in the prediction of Agreeableness, (D) the Pain and emotional processing (**Emo**) networks in the prediction of Neuroticism.

4.4 Predicting Extraversion

Extraversion was predicted by the RSFC patterns within the networks of reward (**Rew**) in Women and face perception (**Face**) in Men. Moreover, in Women, this trait was also significantly predicted by the whole-brain (**Connectome**) RSFC. Extraversion is generally described as behavioural exploration and sensitivity to specific rewards. Importantly, a distinction has been also made between "Agentic Extraversion", reflected in assertiveness, dominance, and ambition aspects, and a "Affiliative Extraversion" which is more related to sociability and affiliative social bonding (DeYoung et al. 2007; c.f. Allen and DeYoung 2016).

As discussed previously in paragraph **4.3**, the traits of Extraversion and Openness exhibit a shared variance, known as "β" factor and are genetically influenced by the dopaminergic system (c.f. Allen and DeYoung 2016). Notably, while for Openness, **Rew**'s most used nodes encompassed the mesocortical pathway (see above), for Extraversion, it was regions along the mesolimbic branch that were mostly used (amygdala, NAc and OFC). Thus, we suggest that even though FC of **Rew** predicts both Openness and Extraversion, the functional connectivity of two different subsystems of the **Rew** network are informative for the two different traits, namely the mesocortical and mesolimbic pathway respectively. In favour of this distinction, extraverts were shown to be more sensitive toward the motivational content of the reward stimulus, encoded by DA neurons along the mesolimbic pathway (Bromberg-Martin et al. 2010; DeYoung 2013). We thus believe that the prediction of Extraversion from the FC within **Rew** might well-capture the "Agentic" dimension of Extraversion, given the motivational value of the rewarding stimuli and drive toward a goal prompted by the dopaminergic mesolimbic system.

While extraversion in Women was found to be associated to FC of **Rew**, relationships of this trait, in Men, were found with FC in **Face** network. Faces are arguably the most important social stimuli for humans and it has been shown that extraverts compared to introvert, by spending more time on people, are significantly better at recognizing faces (Li and Liu 2010). Extraversion's hedonic experience of goal achievement is enclosed in the "Affiliative" component (DeYoung et al. 2007; c.f. Allen and DeYoung 2016) and its genetic variation has been also pointed to the opiate system, due to its involvement in the hedonic response to the stimulus (Peciña et al. 2006). It is therefore possible that the endogenous opioid system via modulation of amygdala and medial prefrontal cortex (Tejeda et al. 2015;

Selleck and Baldo 2017), most used regions in the connections of **Face**, mediate both the perception of faces (Martin et al. 2006) and the social bonding (Pasternak and Pan 2013). We thus suggest that functional connectivity within the **Face** network in Men, is mostly related to the "Affiliative" aspect of Extraversion.

The last prediction of Extraversion is based on whole-brain FC in Women (**Sample 1**: r = 0.29; **Sample 2**: r = 0.23, both p < 0.05; for gender comparison in **Sample 1**, Cohen's q = 0.323, p < 0.05). However, a major issue using whole-brain connectivity patter might be the lack of anatomical localization for the most informative features, as none of them resulted to be used more than 40% of the predictive models, indicating a heterogeneous mosaic of connections which contribute to the prediction of Extraversion. The only theory in personality neuroscience which relates the functioning of entire cortex to Extraversion (and Neuroticism, see below **4.6**) is Eysenck's biological theory of personality (Eysenck 1967). Here, Extraversion is thought to depend on the variability in cortical arousal, with introverted individuals having lower response thresholds consequently more cortical arousal compared to extraverts. In favour of this hypothesis, the topological properties of whole-brain RSFC has shown that brains of more extraverted individuals behave more similarly to a "small-world" compared to a "random" network, with higher clustering coefficient compared to introverts (Gao et al. 2013). A clustered configuration, which supports a more modularized information processing and fault tolerance, can therefore be associated with higher arousal threshold in extraverts' cortex. We also observed that this prediction performance was significantly stronger in Women compared to Men (r = 0.29 in Women and r = -0.03 in Men of **Sample 1**). Again, a possible cause might be the involvement of ovarian hormones, targeting specifically the most densely interconnected hub structures of the connectome (Alawieh et al. 2015) as well as influencing level of Extraversion (Jokela et al. 2009; Ziomkiewicz et al. 2012). However, more studies are needed to prove this interaction between Extraversion, estrogen and the topographical properties of whole-brain functional connectivity.

To sum up, connectivity of regions encoding the motivational value and the drive toward a goal (**Rew**) and the hedonic processing of the goal itself (**Face**), were informative to predict interindividual variability in the trait of Extraversion possibly capturing the "Agentic" and "Affiliative" aspects of the trait respectively (**Fig 4**). Importantly, given the modulation of ovarian hormones on both the trait of Extraversion and on the topological properties of the **Connectome**, we would suggest that sex hormones might be a possible mediator of this trait-network relationship, resulting in better prediction performance in Women.

4.5 Predicting Agreeableness
RSFC patterns in the **AM** network could predict the individual level of perceived Agreeableness while grouping men and women in both samples. This trait reflects a high

desire to avoid interpersonal conflicts (Jensen-Campbell and Graziano 2001) and strong affect regulation (Ryan et al. 2011). In line with this, positive correlations have been demonstrated between Agreeableness and regions supporting social functioning (Hooker et al. 2008; DeYoung et al. 2010; Hassabis et al. 2014) and midline regions of the default mode network (DMN), as deputed to self-referential process (Adelstein et al. 2011; Sampaio et al. 2014). Our prediction of Agreeableness from the **AM** network supports a crucial role of self-reference, strongly linked to autobiographical memory (Molnar-Szakacs and Arzy 2009), in how high agreeable individuals deal with social demands. Self-related cognition has been often discussed at the neural level as the product of interaction between the DMN and the mirror neuron system (MNS), the first responsible for high-level mentalizing function and the second for embodied simulation-based representation (Keysers and Gazzola 2007; Qin and Northoff 2011; c.f. Molnar-Szakacs and Uddin 2013). As a result, the privileged access to the own physical and mental states would allow a better insight into others' physical and mental states, and consequent appropriate social responses.

Interestingly, within the **AM** network, most used connections that informed about the trait in both samples reflected the interaction between the DMN and MNS systems: nodes with highest weights belonged indeed to DMN subsystem, such as medial PFC, posterior cingulate cortex, medial temporal lobe (amygdala and hippocampus) and lateral parietal cortex (temporo-parietal junction). The remaining nodes with the highest weights belonged to the MNS, such as inferior frontal gyrus, precentral gyrus, inferior parietal cortex and superior temporal sulcus. Our result, hence, supports the interplay of these two subsystems in the context of self-processing (here expressed via memory recollection about past experiences, **AM**) and that this knowledge about the self can significantly predict Agreeableness, the trait most reflecting enhanced social skills.

4.6 Predicting Neuroticism

In Men, self-reported Neuroticism was predicted by RSFC within the emotional processing network (**Emo**). Additionally, the RSFC from the whole brain (**Connectome**) significantly predicted this trait across the entire sample and then specifically in Men only. Neuroticism represents a broad dimension of individual differences in the tendency to experience negative, distressing emotions. High Neuroticism scores entail the experience of fear, anger, sadness, embarrassment, the incapacity to control cravings and urges, and to cope with stress (Costa and McCrae 1987). Within this trait, it is possible to delineate two major divisions, one related to the experience of anxiety, fear and passive avoidance, and referred in literature as the component Withdrawal, and the other related to irritability, anger and active defensive responses, or Volatility (DeYoung et al. 2007). Neuroticism is arguably the most studied personality trait and is an important predictor of many different mental and physical disorders (Lahey 2009). Furthermore, the two components of Neuroticism (Withdrawal and Volatility) highly reflect the dimension of Behavioural Inhibition System (BIS) and Fight-

Flight-Freeing System (FFFS) from the Gray's Reinforcement Theory (Gray and Mcnaughton 2000), conceptualized in term of their neurobiology. Interestingly, this distinction between the Volatility/ FFFS and Withdrawal/BIS seems to be captured by the two networks showing predictability power for Neuroticism, **Emo** and **Pain**. Even though this last prediction (**Pain**) was found significant in **Sample 1** (with $r = 0.15$, $p < 0.05$ in Men) but not fully replicated in the **Sample 2** (with $r = 0.2$, $p = 0.05$ in Men) (**Fig 4**), we would still suggest that recruitment of this network in association to Neuroticism might indicate that perception of the aversive stimulus via the **Pain** network (Iannetti and Mouraux 2010; Hayes and Northoff 2012) could lead high-Neuroticism men to inhibit their behaviours such to avoid potential threats and punishments (Withdrawal). Conversely, **Emo** network would trigger emotional responses for either escaping or eliminating the threat, but in both cases showing a strong emotional lability (Volatility). Beyond associations with specific networks, Neuroticism could also be predicted from the whole-brain RSFC (**Connectome**) in Men and across genders. This is nicely in line with graph analysis studies (Gao et al. 2013; Servaas et al. 2015) showing that the neurotic brain displays topological properties of a "random network" and overall weaker FC. Here cortisol might play a specific role, the hormone that is most closely associated with a biological reaction to stress and found to correlate with Neuroticism. However, the directionality of correlation seems to depend on gender: many studies converged in discovering that Neuroticism was positively correlated with baseline cortisol in men, but the opposite was true in women (Zobel et al. 2004; Oswald et al. 2006; DeSoto and Salinas 2015). Thus, especially in men, the overabundance of cortisol by potentiating neuronal degeneration (Sapolsky 1994), might be responsible for the overall smaller brain volume (Liu et al. 2013), white-matter (Bjørnebekk et al. 2013) and gray-matter (Servaas et al. 2015) functional disconnectivity found in high-Neuroticism individuals compared to the more emotional stable. Given that all the three networks (**Emo**, **Pain**, **Connectome**) showed a stronger predictability in Men compared to Women (statistically significant for the first two, and a strong trend for the third, see **Table 3**), we suggest that gender may moderate Neuroticism's relationship to cortisol. However, more (direct) studies are needed to better understand this intricate relationship between RSFC, cortisol, Neuroticism and gender and to shed light on the neural mechanisms that make women's brain more susceptible to Neuroticism-related mental disorders (Jorm 1987).

4.7 Implications for the neurobiology of FFM

Contrary to other important theories of personality, such as Cloninger's Tridimensional Personality Questionnaire (TPQ) or Gray's Reinforcement Sensitivity Theory (RST), the FFM is not based on biological grounds. However, variability in its personality factors had been associated to the brain, given that personality traits are the product of our actions, emotions and, more generally, cognitive processes. In this way, the cognitive mechanisms work as intermediate bridge between the psychometric constructs of personality and plausible biological substrates. However, the relationships among these factors (brain,

behaviour and personality) can be misleading in the context of personality predictions, which, in fact, were significant only to a moderate level, compared to other findings: contrary to predictions of sustain attention (Rosenberg et al. 2016) or reading comprehension (Cui et al. 2017) which tap predictability of cognitive process itself, personality traits are mostly modulators of these cognitive processes. This may make it more difficult to find brain correlates of personality in specific networks associated with those functions.

Also, the hierarchy of the FFM model might have contributed in enlarging the gap: in our findings, we highlighted the possibility that the predictions of one trait from different networks could reflect different components within this trait, also known as facet (cf. DeYoung et al. 2007; Koelsch et al. 2013; Haas et al. 2015). For example, we discussed the prediction of Extraversion from **Rew** and **Face** as potentially capturing the "Agentic" and "Affiliative" aspects respectively, or the prediction of Neuroticism from **Pain** and **Emo** as linked to Withdrawal and Volatility. Conversely, when the same network was predicting two different traits (e.g. **Rew** predicting Openness and Extraversion, discussed in light of the saliency and motivational contribution for the two traits), the prediction might have indeed boosted if investigating the meta-trait "β", which reflects their shared variance within the dopaminergic system and thus more prone to be predicted by the network of reward processing (DeYoung 2013). Therefore, the level of abstraction of the five traits might not mapped well to particular brain systems, and more studies are encouraged for testing both more specific and homogeneous sub-dimensions as well as more heterogeneous higher-order factor structure. Lastly, many biological mechanisms participate in evoking the same cognitive process, e.g. changes in brain structure, function, or genetic, which are then intrinsically connected with personality. We here used RSFC as "marker" for the individual expression of personality traits, enduring across time and situations. However, a downside of FC in resting conditions might be that it has not so much to do with how personality factors come together to "produce" stable modulations of a whole range of cognitive processes. Therefore, other brain measurements (as structural connectivity, task-based functional activation, or molecular genetics) might be also useful in gaining more knowledge on the biology of personality and its relationship with specific mental functions. Keeping in mind that we cannot expect biological mechanisms to show clear-cut as the respective psychometric dimensions (Yarkoni 2015), but conversely many biological mechanisms (function, structure, neurotransmitters) as well as many mental functions can be informative for a given personality trait, we therefore support the need for a multi-level approach in future studies as proposed by Yarkoni in order to achieve a unified description of the biological bases of personality traits.

However, even though all these aspects might affect the relationship between brain function (and structure) and personality, we here do provide insights on the relation between brain and personality: when analysing the entire sample while adjusting for gender effects, only two

predictions (**VA** predicting Openness and **AM** predicting Agreeableness) can be found not specifically driven by one gender-group. However, when looking at men and women separately, we observed much more and larger effects, evidence which highly remarks the importance of gender while investigating the neural correlates of personality. Specifically, the current findings propose a link between Openness and executive and affective domain. Agreeableness with memory domain. Extraversion with social and affective networks and lastly Neuroticism with the affective system. Interestingly, these last two traits could be predicted as well from the entire **Connectome**. An interesting consideration is that Openness could be significantly predicted by three different, barely overlapping networks (**Pain**, **Rew**, **VA**), but could not be predicted from the whole-brain, which was covering the nodes of all the three at the same time. We thus argue for a better predictability of Openness from specific and separate functional networks. Contrarily, Extraversion and Neuroticism could be significantly predicted by both meta-analytic networks and the whole-brain, pointing to the importance of also global effects, besides specific functions. This is particularly true for Extraversion, which showed significantly higher prediction performance from global RSFC (**Connectome**) with a very vast nodes contribution, rather than from the specific networks of **Rew** and **Face**, thus favouring the global effects over the specific functions for this trait.

4.8 Conclusions

Using multivariate machine learning, we showed that personality traits can be predicted from RSFC patterns in affective, social, executive and memory networks of the brain, as well as from the whole-brain. Our observation that for most of these networks predictive power was gender-specific complements previous morphometric findings (Nostro et al. 2016) in highlighting the crucial role of gender when trying to understand the neurobiology of personality. Additionally, the many-to-many associations between mental functions and personality traits, indicate the complexity of the biological substrates of personality, as many functional systems may contribute to the observable differences in each trait (for a critical review see Yarkoni 2015). Maybe even more fundamental are the implications for the concept of personality, given that even a trait as complex and broad as, for instance, Openness, seems to have a neurobiological underpinning in pre-defined functional networks that enables estimation of the individual level of that trait in a new subject.

References

Adelstein JS, Shehzad Z, Mennes M, et al (2011) Personality Is Reflected in the Brain's Intrinsic Functional Architecture. PLoS One 6:e27633. doi: 10.1371/journal.pone.0027633

Alawieh A, Sabra Z, Sabra M, et al (2015) A rich-club organization in brain ischemia protein interaction network. Sci Rep. doi: 10.1038/srep13513

Allen EA, Erhardt EB, Damaraju E, et al (2011) A baseline for the multivariate comparison of resting-state networks. Front Syst Neurosci 5:2. doi: 10.3389/fnsys.2011.00002

Allen TA, DeYoung CG (2016) Personality Neuroscience and the Five Factor Model.

Arélin K, Mueller K, Barth C, et al (2015) Progesterone mediates brain functional connectivity changes during the menstrual cycle-a pilot resting state MRI study. Front Neurosci 9:1–11. doi: 10.3389/fnins.2015.00044

Asahi S, Okamoto Y, Okada G, et al (2004) Negative correlation between right prefrontal activity during response inhibition and impulsiveness: A fMRI study. Eur Arch Psychiatry Clin Neurosci 254:245–251. doi: 10.1007/s00406-004-0488-z

Ashburner J, Friston KJ (2005) Unified segmentation. Neuroimage 26:839–851. doi: 10.1016/j.neuroimage.2005.02.018

Bachrach Y, Kosinski M, Graepel T, et al (2012) Personality and patterns of Facebook usage. Proc 3rd Annu ACM Web Sci Conf - WebSci '12 24–32. doi: 10.1145/2380718.2380722

Baumgartner U, Buchholz HG, Bellosevich A, et al (2006) High opiate receptor binding potential in the human lateral pain system. Neuroimage 30:692–699. doi: 10.1016/j.neuroimage.2005.10.033

Beaty RE, Kaufman SB, Benedek M, et al (2016) Personality and complex brain networks: The role of openness to experience in default network efficiency. Hum Brain Mapp 37:773–779. doi: 10.1002/hbm.23065

Berridge KC, Robinson TE (1998) What is the role of dopamine in reward: Hedonic impact, reward learning, or incentive salience? Brain Res Rev 28:309–369. doi: 10.1016/S0165-0173(98)00019-8

Binder JR, Desai RH, Graves WW, Conant LL (2009) Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. Cereb Cortex 19:2767–2796. doi: 10.1093/cercor/bhp055

Bjørnebekk A, Fjell AM, Walhovd KB, et al (2013) Neuronal correlates of the five factor model (FFM) of human personality:Multimodal imaging in a large healthy sample. Neuroimage 65:194–208. doi: 10.1016/j.neuroimage.2012.10.009

Bromberg-Martin ES, Matsumoto M, Hikosaka O (2010) Dopamine in Motivational Control: Rewarding, Aversive, and Alerting. Neuron 68:815–834.

Bzdok D, Schilbach L, Vogeley K, et al (2012) Parsing the neural correlates of moral cognition: ALE meta-analysis on morality, theory of mind, and empathy. Brain Struct Funct

217:783–796. doi: 10.1007/s00429-012-0380-y

Ciric R, Wolf DH, Power JD, et al (2017) Benchmarking of participant-level confound regression strategies for the control of motion artifact in studies of functional connectivity. Neuroimage 154:174–187. doi: 10.1016/j.neuroimage.2017.03.020

Cohen J (1988) Statistical power analysis for the behavioral sciences. Stat. Power Anal. Behav. Sci. 2nd:567.

Costa PT, McCrae RR (1992) Professional manual: revised NEO personality inventory (NEO-PI-R) and NEO five-factor inventory (NEO-FFI).

Costa PT, McCrae RR (1987) Neuroticism, Somatic Complaints, and Disease: Is the Bark Worse than the Bite? J Pers 55:299–316. doi: 10.1111/j.1467-6494.1987.tb00438.x

Cremers HR, Demenescu LR, Aleman A, et al (2010) Neuroticism modulates amygdala-prefrontal connectivity in response to negative emotional facial expressions. Neuroimage 49:963–970. doi: 10.1016/j.neuroimage.2009.08.023

Cui Z, Su M, Li L, et al (2017) Individualized Prediction of Reading Comprehension Ability Using Gray Matter Volume Individualized Prediction of Reading Comprehension Ability Using Gray Matter Volume. doi: 10.1093/cercor/bhx061

De Vico Fallani F, Richiardi J, Chavez M, Achard S (2014) Graph analysis of functional brain networks: practical issues in translational neuroscience. Philos Trans R Soc B Biol Sci 369:20130521–20130521. doi: 10.1098/rstb.2013.0521

Denkova E, Dolcos S, Dolcos F (2012) Reliving Emotional Personal Memories: Affective Biases Linked to Personality and Sex-Related Differences. Emotion 12:515–528. doi: 10.1037/a0026809.supp

DeSoto MC, Salinas M (2015) Neuroticism and cortisol: The importance of checking for sex differences. Psychoneuroendocrinology 62:174–179. doi: 10.1016/j.psyneuen.2015.07.608

DeYoung C (2014) Openness/Intellect: a dimension of personality reflecting cognitive exploration. APA Handb Personal Soc Psychol Personal Process Individ Differ 4:369–399. doi: dx.doi.org/10.1037/14343-017

DeYoung CG (2015) Cybernetic Big Five Theory. J Res Pers 56:33–58. doi: 10.1016/j.jrp.2014.07.004

DeYoung CG (2010) Personality Neuroscience and the Biology of Traits. Soc Personal Psychol Compass 4:1165–1180. doi: 10.1111/j.1751-9004.2010.00327.x

DeYoung CG (2013) The neuromodulator of exploration: A unifying theory of the role of dopamine in personality. Front Hum Neurosci. doi: 10.3389/fnhum.2013.00762

DeYoung CG, Gray JR (2009) Personality neuroscience: explaining individual differences in affect, behaviour and cognition.

DeYoung CG, Hirsh JB, Shane MS, Papademetris X (2010) Testing Predictions From Personality Neuroscience: Brain Structure and the Big Five. 21:820–828. doi: 10.1177/0956797610370159.Testing

DeYoung CG, Peterson JB, Higgins DM (2005) Sources of Openness/Intellect: Cognitive and neuropsychological correlates of the fifth factor of personality. J Pers 73:825–858. doi:

10.1111/j.1467-6494.2005.00330.x

DeYoung CG, Quilty LC, Peterson JB (2007) Between facets and domains: 10 aspects of the Big Five. J Pers Soc Psychol 93:880–896. doi: 10.1037/0022-3514.93.5.880

Doyle OM, Mehta MA, Brammer MJ (2015) The role of machine learning in neuroimaging for drug discovery and development. Psychopharmacology (Berl) 232:4179–4189. doi: 10.1007/s00213-015-3968-0

Dreher J-C, Schmidt PJ, Kohn P, et al (2007) Menstrual cycle phase modulates reward-related neural function in women. Proc Natl Acad Sci U S A 104:2465–70. doi: 10.1073/pnas.0605569104

Dubois J, Galdi P, Han Y, et al (2017) Predicting personality traits from resting-state fMRI . bioRxiv. doi: 10.1101/215129

Eickhoff SB, Grefkes C (2011) Approaches for the Integrated Analysis of Structure, Function and Connectivity of the Human Brain. Clin EEG Neurosci 42:107–121. doi: 10.1177/155005941104200211

Eickhoff SB, Paus T, Caspers S, et al (2007) Assignment of functional activations to probabilistic cytoarchitectonic areas revisited. Neuroimage 36:511–21. doi: 10.1016/j.neuroimage.2007.03.060

Eickhoff SB, Stephan KE, Mohlberg H, et al (2005) A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. Neuroimage 25:1325–1335. doi: 10.1016/j.neuroimage.2004.12.034

Eysenck HJ (1967) Biological basis of personality. Nature 199:1031–1034. doi: 10.1038/1991031a0

Feinberg DA, Moeller S, Smith SM, et al (2010) Multiplexed echo planar imaging for sub-second whole brain fmri and fast diffusion imaging. PLoS One. doi: 10.1371/journal.pone.0015710

Filippi M, Valsasina P, Misci P, et al (2013) The organization of intrinsic brain activity differs between genders: A resting-state fMRI study in a large cohort of young healthy subjects. Hum Brain Mapp 34:1330–1343. doi: 10.1002/hbm.21514

Fleischhauer M, Enge S, Brocke B, et al (2010) Same or different? Clarifying the relationship of need for cognition to personality and intelligence. Personal Soc Psychol Bull 36:82–96. doi: 10.1177/0146167209351886

Fox PT, Lancaster JL, Laird AR, Eickhoff SB (2014) Meta-Analysis in Human Neuroimaging: Computational Modeling of Large-Scale Databases Peter. Annu Rev Neurosci 37:409–434. doi: 10.1002/aur.1474.Replication

Gao Q, Xu Q, Duan X, et al (2013) Extraversion and neuroticism relate to topological properties of resting-state brain networks. Front Hum Neurosci 7:257. doi: 10.3389/fnhum.2013.00257

Gazzola V, Aziz-Zadeh L, Keysers C (2006) Empathy and the Somatotopic Auditory Mirror System in Humans. Curr Biol 16:1824–1829. doi: 10.1016/j.cub.2006.07.072

Glasser MF, Sotiropoulos SN, Wilson JA, et al (2013) The minimal preprocessing pipelines

for the Human Connectome Project. Neuroimage 80:105–124. doi: 10.1016/j.neuroimage.2013.04.127

Golbeck J (2011) Predicting Personality with Social Media. Proc 2011 Annu Conf Ext Abstr Hum factors Comput Syst CHI EA 11 253–262. doi: 10.1145/1979742.1979614

Golbeck J, Robles C, Edmondson M, Turner K (2011) Predicting personality from twitter. Proc - 2011 IEEE Int Conf Privacy, Secur Risk Trust IEEE Int Conf Soc Comput PASSAT/SocialCom 2011 149–156. doi: 10.1109/PASSAT/SocialCom.2011.33

Gray J a, Mcnaughton N (2000) The Neuropsychology of Anxiety: An Enquiry into the Functions of the Septo-Hippocampal System, Second. Oxford Psychol Ser Second Edi:433. doi: 10.1017/S0140525X00013066

Grosbras MH, Beaton S, Eickhoff SB (2012) Brain regions involved in human movement perception: A quantitative voxel-based meta-analysis. Hum Brain Mapp 33:431–454. doi: 10.1002/hbm.21222

Haas BW, Brook M, Remillard L, et al (2015) I know how you feel: The warm-altruistic personality profile and the empathic brain. PLoS One 10:1–15. doi: 10.1371/journal.pone.0120639

Harrison PJ, Tunbridge EM (2008) Catechol-O-methyltransferase (COMT): a gene contributing to sex differences in brain function, and to sexual dimorphism in the predisposition to psychiatric disorders. Neuropsychopharmacology 33:3037–3045. doi: 10.1038/sj.npp.1301543

Hassabis D, Spreng RN, Rusu AA, et al (2014) Imagine all the people: How the brain creates and uses personality models to predict behavior. Cereb Cortex 24:1979–1987. doi: 10.1093/cercor/bht042

Hastie T, Tibshirani R, Friedman J (2009) The Elements of Statistical Learning. Elements 1:337–387. doi: 10.1007/b94608

Hayes DJ, Northoff G (2012) Common brain activations for painful and non-painful aversive stimuli. BMC Neurosci 13:60. doi: 10.1186/1471-2202-13-60

Hjelmervik H, Hausmann M, Osnes B, et al (2014) Resting states are resting traits - An fMRI study of sex differences and menstrual cycle effects in resting state cognitive control networks. PLoS One 9:32–36. doi: 10.1371/journal.pone.0103492

Holmes CJ, Hoge R, Collins L, et al (1998) Enhancement of MR images using registration for signal averaging. J Comput Assist Tomogr 22:324–333. doi: 10.1097/00004728-199803000-00032

Hooker CI, Verosky SC, Miyakawa A, et al (2008) The influence of personality on neural mechanisms of observational fear and reward learning. Neuropsychologia 46:2709–2724. doi: 10.1016/j.neuropsychologia.2008.05.005

Hu X, Erb M, Ackermann H, et al (2011) Voxel-based morphometry studies of personality: Issue of statistical model specification—effect of nuisance covariates. Neuroimage 54:1994–2005. doi: 10.1016/j.neuroimage.2010.10.024

Iacoboni M (2009) Imitation, Empathy, and Mirror Neurons. Annu Rev Psychol 60:653–670.

doi: 10.1146/annurev.psych.60.110707.163604

Iannetti GD, Mouraux A (2010) From the neuromatrix to the pain matrix (and back). Exp. Brain Res. 205:1–12.

IBM Corp. Released (2011) IBM SPSS Statistics for Windows, Version 20.0.

Jensen-Campbell LA, Graziano WG (2001) Agreeableness as a moderator of interpersonal conflict. J Pers 69:323–361. doi: 10.1111/1467-6494.00148

Jokela M, Kivimäki M, Elovainio M, Keltikangas-Järvinen L (2009) Personality and having children: A two-way relationship. J Pers Soc Psychol 96:218–230. doi: 10.1037/a0014058

Jorm AF (1987) Sex differences in neuroticism: A quantitative synthesis of published research. Aust N Z J Psychiatry 21:501–506. doi: 10.3109/00048678709158917

Kennis M, Rademaker AR, Geuze E (2013) Neural correlates of personality: An integrative review. Neurosci Biobehav Rev 37:73–95. doi: 10.1016/j.neubiorev.2012.10.012

Keysers C, Gazzola V (2007) Integrating simulation and theory of mind: from self to social cognition. Trends Cogn. Sci. 11:194–196.

Kim H, Shimojo S, O'Doherty JP (2006) Is avoiding an aversive outcome rewarding? Neural substrates of avoidance learning in the human brain. PLoS Biol 4:1453–1461. doi: 10.1371/journal.pbio.0040233

Koelsch S, Skouras S, Jentschke S (2013) Neural correlates of emotional personality: A structural and functional magnetic resonance imaging study. PLoS One. doi: 10.1371/journal.pone.0077196

Kogler L, Müller VI, Chang A, et al (2015) Psychosocial versus physiological stress - Meta-analyses on deactivations and activations of the neural correlates of stress reactions. Neuroimage 119:235–251. doi: 10.1016/j.neuroimage.2015.06.059

Konishi Y, Tanii H, Otowa T, et al (2014) Gender-specific association between the COMT Val158Met polymorphism and openness to experience in panic disorder patients. Neuropsychobiology 69:165–174. doi: 10.1159/000360737

Kosinski M, Stillwell D, Graepel T (2013) Private traits and attributes are predictable from digital records of human behavior. Proc Natl Acad Sci U S A 110:5802–5. doi: 10.1073/pnas.1218772110

Kumari V (2004) Personality Predicts Brain Responses to Cognitive Demands. J Neurosci 24:10636–10641. doi: 10.1523/JNEUROSCI.3206-04.2004

Kunisato Y, Okamoto Y, Okada G, et al (2011) Personality traits and the amplitude of spontaneous low-frequency oscillations during resting state. Neurosci Lett 492:109–113. doi: 10.1016/j.neulet.2011.01.067

Lahey BB (2009) Public health significance of neuroticism. Am Psychol 64:241–256. doi: 10.1037/a0015309

Lane DM (2013) Introduction To Statistics. Introd to Stat 454–8. doi: 10.1016/B978-0-12-370483-2.00006-0

Langner R, Eickhoff SB (2013) Sustaining attention to simple tasks: a meta-analytic review of the neural mechanisms of vigilant attention. Psychol Bull 139:870–900. doi:

10.1037/a0030694

Lee M, Smyser C, Shimony J (2012) {Resting-State} {fMRI:} A Review of Methods and Clinical Applications. {AJNR} Am J Neuroradiol. doi: 10.3174/ajnr.A3263

Lei X, Yang T, Wu T (2015) Functional neuroimaging of extraversion-introversion. Neurosci Bull 31:663–675. doi: 10.1007/s12264-015-1565-1

Li J, Liu J (2010) Extraversion predicts individual differences in face recognition. 295–298.

Li N, Ma N, Liu Y, et al (2013) Resting-State Functional Connectivity Predicts Impulsivity in Economic Decision-Making. J Neurosci 33:4886–4895. doi: 10.1523/JNEUROSCI.1342-12.2013

Liu W-Y, Weber B, Reuter M, et al (2013) The Big Five of Personality and structural imaging revisited: a VBM - DARTEL study. Neuroreport 24:375–80. doi: 10.1097/WNR.0b013e328360dad7

Liu X, Hairston J, Schrier M, Fan J (2011) Common and distinct networks underlying reward valence and processing stages: A meta-analysis of functional neuroimaging studies. Neurosci Biobehav Rev 35:1219–1236. doi: 10.1016/j.neubiorev.2010.12.012

MacLean MH, Arnell KM (2010) Personality predicts temporal attention costs in the attentional blink paradigm. Psychon Bull Rev 17:556–562. doi: 10.3758/PBR.17.4.556

Madsen MK, Mc Mahon B, Andersen SB, et al (2015) Threat-related amygdala functional connectivity is associated with 5-HTTLPR genotype and neuroticism. Soc Cogn Affect Neurosci 11:140–149. doi: 10.1093/scan/nsv098

Magyar M, Gonda X, Pap D, et al (2017) Decreased openness to experience is associated with migraine-type headaches in subjects with lifetime depression. Front Neurol. doi: 10.3389/fneur.2017.00270

Marcus B, Machilek F, Schütz A (2006) Personality in cyberspace: personal Web sites as media for personality expressions and impressions. J Pers Soc Psychol 90:1014–1031. doi: 10.1037/0022-3514.90.6.1014

Martin L, Clair J, Davis P, et al (2006) Enhanced recognition of facial expressions of disgust in opiate users receiving maintenance treatment. Addiction 101:1598–1605. doi: 10.1111/j.1360-0443.2006.01574.x

McCrae RR, Costa PT (2004) A contemplated revision of the NEO Five-Factor Inventory. Pers Individ Dif 36:587–596. doi: 10.1016/S0191-8869(03)00118-1

Moeller S, Yacoub E, Olman CA, et al (2010) Multiband multislice GE-EPI at 7 tesla, with 16-fold acceleration using partial parallel imaging with application to high spatial and temporal whole-brain FMRI. Magn Reson Med 63:1144–1153. doi: 10.1002/mrm.22361

Mohan G, Mulla ZR (2013) OPENNESS TO EXPERIENCE AND WORK OUTCOMES: EXPLORING THE MODERATING EFFECTS OF CONSCIENTIOUSNESS AND JOB COMPLEXITY.

Molnar-Szakacs I, Arzy S (2009) Searching for an integrated self-representation. Commun. Integr. Biol. 2:365–367.

Molnar-Szakacs I, Uddin LQ (2013) Self-Processing and the Default Mode Network:

Interactions with the Mirror Neuron System. Front Hum Neurosci. doi: 10.3389/fnhum.2013.00571

Nostro AD, Müller VI, Reid AT, Eickhoff SB (2016) Correlations Between Personality and Brain Structure: A Crucial Role of Gender. Cereb Cortex 1–15. doi: 10.1093/cercor/bhw191

Oktar N, Oktar Y (2015) Machine Learning and Neuroimaging. J Neurol Sci [Turkish] 32:1–4.

Orrù G, Pettersson-Yeo W, Marquand AF, et al (2012) Using Support Vector Machine to identify imaging biomarkers of neurological and psychiatric disease: A critical review. Neurosci Biobehav Rev 36:1140–1152. doi: 10.1016/j.neubiorev.2012.01.004

Oswald LM, Zandi P, Nestadt G, et al (2006) Relationship between Cortisol Responses to Stress and Personality. 1583–1591. doi: 10.1038/sj.npp.1301012

Pang Y, Cui Q, Wang Y, et al (2016) Extraversion and neuroticism related to the resting-state effective connectivity of amygdala. Sci Rep. doi: 10.1038/srep35484

Passamonti L, Fera F, Magariello A, et al (2006) Monoamine oxidase-A genetic variations influence brain activity associated with inhibitory control: New insight into the neural correlates of impulsivity. Biol Psychiatry 59:334–340. doi: 10.1016/j.biopsych.2005.07.027

Passamonti L, Terracciano A, Riccelli R, et al (2015) Increased functional connectivity within mesocortical networks in open people. Neuroimage 104:301–309. doi: 10.1016/j.neuroimage.2014.09.017

Pasternak GW, Pan Y-X (2013) Mu opioids and their receptors: evolution of a concept. Pharmacol Rev 65:1257–317. doi: 10.1124/pr.112.007138

Pearman A (2009) Predictors of Subjective Memory in Young Adults. 101–107. doi: 10.1007/s10804-009-9063-1

Peciña S, Smith KS, Berridge KC (2006) Hedonic Hot Spots in the Brain. Neurosci 12:500–511. doi: 10.1177/1073858406293154

Plitt M, Barnes KA, Wallace GL, et al (2015) Resting-state functional connectivity predicts longitudinal change in autistic traits and adaptive functioning in autism. Proc Natl Acad Sci 112:E6699-6706. doi: 10.1073/pnas.1510098112

Power JD, Cohen AL, Nelson SM, et al (2011) Functional network organization of the human brain. Neuron 72:665–678. doi: 10.1016/j.neuron.2011.09.006.Functional

Qin P, Northoff G (2011) How is our self related to midline regions and the default-mode network? Neuroimage 57:1221–1233.

Quercia D, Kosinski M, Stillwell D, Crowcroft J Our Twitter Profiles , Our Selves : Predicting Personality with Twitter.

Rosenberg MD, Finn ES, Scheinost D, et al (2016) A neuromarker of sustained attention from whole-brain functional connectivity. Nat Neurosci 19:165–71. doi: 10.1038/nn.4179

Rottschy C, Langner R, Dogan I, et al (2012) Modelling neural correlates of working memory: A coordinate-based meta-analysis. Neuroimage 60:830–846. doi: 10.1016/j.neuroimage.2011.11.050

Ryan JP, Sheu LK, Gianaros PJ (2011) Resting state functional connectivity within the

cingulate cortex jointly predicts agreeableness and stressor-evoked cardiovascular reactivity. Neuroimage 55:363–370. doi: 10.1016/j.neuroimage.2010.11.064

Sabatinelli D, Fortune EE, Li Q, et al (2011) Emotional perception: Meta-analyses of face and natural scene processing. Neuroimage 54:2524–2533. doi: 10.1016/j.neuroimage.2010.10.011

Salimi-khorshidi G, Douaud G, Beckmann CF, et al (2014) Automatic denoising of functional MRI data : Combining independent component analysis and hierarchical fusion of classi fi ers. Neuroimage 90:449–468. doi: http://dx.doi.org/10.1016/j.neuroimage.2013.11.046

Salimi-Khorshidi G, Douaud G, Beckmann CF, et al (2014) Automatic denoising of functional MRI data: Combining independent component analysis and hierarchical fusion of classifiers. Neuroimage 90:449–468. doi: 10.1016/j.neuroimage.2013.11.046

Samartsidis P, Montagna S, Nichols TE, Johnson TD (2017) The coordinate-based meta-analysis of neuroimaging data. Stat Sci Volume 32:580–599.

Sampaio A, Soares JM, Coutinho J, et al (2014) The Big Five default brain: functional evidence. Brain Struct Funct 219:1913–1922. doi: 10.1007/s00429-013-0610-y

Sapolsky RM (1994) Glucocorticoids, stress and exacerbation of excitotoxic neuron death. Semin. Neurosci. 6:323–331.

Satterthwaite TD, Elliott MA, Gerraty RT, et al (2013) An improved framework for confound regression and filtering for control of motion artifact in the preprocessing of resting-state functional connectivity data. Neuroimage 64:240–256. doi: 10.1016/j.neuroimage.2012.08.052

Schilbach L, Derntl B, Aleman A, et al (2016) Differential Patterns of Dysconnectivity in Mirror Neuron and Mentalizing Networks in Schizophrenia. Schizophr Bull sbw015. doi: 10.1093/schbul/sbw015

Selleck RA, Baldo BA (2017) Feeding-modulatory effects of mu-opioids in the medial prefrontal cortex: a review of recent findings and comparison to opioid actions in the nucleus accumbens. Psychopharmacology (Berl). 234:1439–1449.

Servaas MN, Geerligs L, Renken RJ, et al (2015) Connectomics and neuroticism: an altered functional network organization. Neuropsychopharmacology 40:296–304. doi: 10.1038/npp.2014.169

Shanmugan S, Epperson CN (2014) Estrogen and the prefrontal cortex: Towards a new understanding of estrogen's effects on executive functions in the menopause transition. Hum. Brain Mapp. 35:847–865.

Smith SM, Miller KL, Salimi-Khorshidi G, et al (2011) Network modelling methods for FMRI. Neuroimage 54:875–891. doi: 10.1016/j.neuroimage.2010.08.063

Spreng RN, Mar RA, Kim ASN (2008) The Common Neural Basis of Autobiographical Memory, Prospection, Navigation, Theory of Mind, and the Default Mode: A Quantitative Meta-analysis. J Cogn Neurosci 21:489–510. doi: 10.1162/jocn.2008.21029

Sprouse-Blum AS, Smith G, Sugai D, Parsa FD (2010) Understanding endorphins and their

importance in pain management. Hawaii Med J 69:70–1.

Studer-Luethi B, Jaeggi SM, Buschkuehl M, Perrig WJ (2012) Influence of neuroticism and conscientiousness on working memory training outcome. Pers Individ Dif 53:44–49. doi: 10.1016/j.paid.2012.02.012

Sutin AR, Terracciano A, Kitner-Triolo MH, et al (2011) Personality traits prospectively predict verbal fluency in a lifespan sample. Psychol Aging 26:994–999. doi: 10.1037/a0024276

Tejeda HA, Hanks AN, Scott L, et al (2015) Prefrontal Cortical Kappa Opioid Receptors Attenuate Responses to Amygdala Inputs. Neuropsychopharmacology 40:2856–2864. doi: 10.1038/npp.2015.138

Tipping M (2001) Sparse Bayesian Learning and the Relevance Vector Mach. J Mach Learn Res 1:211–244. doi: 10.1162/15324430152748236

Tzschoppe J, Nees F, Banaschewski T, et al (2014) Aversive learning in adolescents: Modulation by amygdala-prefrontal and amygdala-hippocampal connectivity and neuroticism. Neuropsychopharmacology 39:875–884. doi: 10.1038/npp.2013.287

Van Essen DC, Smith SM, Barch DM, et al (2013) The WU-Minn Human Connectome Project: An overview. Neuroimage 80:62–79. doi: 10.1016/j.neuroimage.2013.05.041

Varikuti DP, Hoffstaedter F, Genon S, et al (2016) Resting-state test retest reliability of a priori defined canonical networks over different preprocessing steps. Brain Struct Funct 1–22. doi: 10.1007/s00429-016-1286-x

Varoquaux G, Raamana P, Engemann D, et al (2016) Assessing and tuning brain decoders: cross-validation, caveats, and guidelines. arXiv:160605201 [statML] 1–14. doi: 10.1016/j.neuroimage.2016.10.038

Varoquaux G, Thirion B (2014) How machine learning is shaping cognitive neuroimaging. Gigascience 3:28. doi: 10.1186/2047-217X-3-28

Viken RJ, Rose RJ, Kaprio J, Koskenvuo M (1994) A developmental genetic analysis of adult personality: Extraversion and neuroticism from 18 to 59 years of age. J Pers Soc Psychol 66:722–730. doi: 10.1037//0022-3514.66.4.722

Wang Y, Fan Y, Bhatt P, Davatzikos C (2010) High-dimensional pattern regression using machine learning: From medical images to continuous clinical variables. Neuroimage 50:1519–1535. doi: 10.1016/j.neuroimage.2009.12.092

Weis S, Hodgetts S, Hausmann M (2017) Sex differences and menstrual cycle effects in cognitive and sensory resting state networks. Brain Cogn.

Xia M, Wang J, He Y (2013) BrainNet Viewer: A Network Visualization Tool for Human Brain Connectomics. PLoS One. doi: 10.1371/journal.pone.0068910

Xu J, Moeller S, Auerbach EJ, et al (2013) Evaluation of slice accelerations using multiband echo planar imaging at 3T. Neuroimage 83:991–1001. doi: 10.1016/j.neuroimage.2013.07.055

Yadollahi P, Khalaginia Z, Vedadhir A, et al (2014) The study of predicting role of personality traits in the perception of labor pain. Iran J Nurs Midwifery Res 19:S97–S102.

Yang W, Cun L, Du X, et al (2015) Gender differences in brain structure and resting-state functional connectivity related to narcissistic personality. Sci Rep 5:10924. doi: 10.1038/srep10924

Yarkoni T (2015) Neurobiological substrates of personality: A critical overview. APA Handb Personal Soc Psychol Vol 4 Personal Process Individ Differ 4:61–83. doi: 10.1017/CBO9781107415324.004

Ziomkiewicz A, Wichary S, Bochenek D, et al (2012) Temperament and ovarian reproductive hormones in women: Evidence from a study during the entire menstrual cycle. Horm Behav 61:535–540. doi: 10.1016/j.yhbeh.2012.01.017

Zobel A, Barkow K, Schulze-Rauschenbach S, et al (2004) High neuroticism and depressive temperament are associated with dysfunctional regulation of the hypothalamic-pituitary-adrenocortical system in healthy volunteers. Acta Psychiatr Scand 109:392–399. doi: 10.1111/j.1600-0447.2004.00313.x

Supplementary material

<u>Table S1: Influence of zygosity on the traits distribution</u>

We performed a Kolmogorov-Smirnov (KS) test in order to verify that the distribution for each trait in monozygotic and dizygotic twins was not significantly different (null hypothesis). Therefore, from the S1200 release we selected only twin participants (N= 563) and later extracted a subsample of unrelated subjects (N = 262, 131 males and 131 females). All the statistics result not significant, i.e. the distribution of each trait in Mz and Dz does not differ.

Trait	K-S statistic (Mz vs Dz)	P value
Openness	0.10	0.47
Conscientiousness	0.06	0.96
Extraversion	0.07	0.87
Agreeableness	0.13	0.23
Neuroticism	0.07	0.93

<u>Table S2: Correlations between factors</u>

Supplementary Table 1: Intercorrelations (Pearson's r) among the 5 personality factors for Sample 1 and Sample 2, across the overall samples, in males, and females.

Sample 1

		Openness	Conscientiousness	Extraversion	Agreeableness	Neuroticism
Openness	**Overall**	-	-0.14*/	0.07/	0.17*/	0.0/
	Males		-0.15/	0.06/	0.17*/	0.07/
	Females		-0.11	0.09	0.18*	-0.08
Conscientiousness	**Overall**	-	-	0.27*/	0.19*/	-0.35*/
	Males			0.32*/	0.24*/	-0.37*/
	Females			0.24*	0.12	-0.36*
Extraversion	**Overall**	-	-	-	0.26*/	-0.32*/
	Males				0.23*/	-0.32*/
	Females				0.34*	-0.3*
Agreeableness	**Overall**	-	-	-	-	-0.26*/
	Males					-0.29*/
	Females					-0.31*
Neuroticism		-	-	-	-	-

Sample 2

		Openness	Conscientiousness	Extraversion	Agreeableness	Neuroticism
Openness	**Overall**	-	-0.17*/	0.13/	0.13/	0.07/
	Males		-0.11/	0.09/	0.13/	0.09/
	Females		-0.2	0.18	0.18	0.08

Conscientious ness	**Overall**	-	-	0.25*/	0.21*/	-0.47*/
	Males			0.32*/	0.26*/	-0.54*/
	Females			0.17	0.13	-0.43*
Extraversion	**Overall**	-	-	-	0.43*/	-0.41*/
	Males				0.40*/	-0.42*/
	Females				0.46*	-0.41*
Agreeableness	**Overall**	-	-	-	-	-0.39*/
	Males					-0.39*/
	Females					-0.45*
Neuroticism		-	-	-	-	-

* Marks significance at p<0.05 (Bonferroni corrected)

Table S3: Coordinates of each network included in the RS functional connectivity network analysis

Empathy Bzdok et al., 2012					
x	**y**	**z**	**Macroanatomical location**	**Original labeling in the Meta-analysis**	**Cytoarchitectonic Assignment**
2.0	56.0	18.0	rdmPFC	dmPFC	Area p32
-8.0	54.0	34.0	ldmPFC	dmPFC	-
36.0	22.0	-8.0	raIns/IFG	raIns	-
54.0	16.0	20.0	rIFG	rIFG	Area45
50.0	30.0	4.0	rIFG (p.Tr)	rIFG	-
-30.0	20.0	4.0	laIns	laIns	-
50.0	12.0	-8.0	rSTG	rIFG	-
-44.0	24.0	-6.0	lIFG(p.Orb)	lIFG	-
-4.0	18.0	50.0	SMA	SMA	
-2.0	28.0	20.0	aMCC	aMCC	Area 33
-4.0	42.0	18.0	pACC	rostral ACC	Areap32
-2.0	-32.0	28.0	PCC	PCC	Retrosplenial Area a30
52.0	-58.0	22.0	rTPJ	rTPJ	Area PGp
-56.0	-58.0	22.0	lTPJ	lTPJ	Area PGa
22.0	-2.0	-16.0	rAm	rAm	Amygdala: SF, CM
54.0	-8.0	-16.0	rMTG	rMTG	-
52.0	-36.0	2.0	rpSTS	rpSTS	-
-12.0	-4.0	12.0	laTh	laTh	Th:Prefrontal,

x	y	z	Macroanatomical Location	Original labeling in the Meta-analysis	Cytoarchitectonic Assignment
6.0	-32.0	2.0	rpTh	rpTh	
26.0	-26.0	-12.0	r Hippo	rHippo	Subiculum
2.0	-20.0	-12.0	Midbrain	Midbrain	-
14.0	4.0	0.0	rGP	rGP	Th:Prefrontal

<table>
<tr><td colspan="6" align="center">Face processing
Grosbras et al., 2012</td></tr>
</table>

x	y	z	Macroanatomical Location	Original labeling in the Meta-analysis	Cytoarchitectonic Assignment
42.0	-78.0	-8.0	r lOcC	r lOcC	hOc4la
-40.0	-82.0	-8.0	lOcC	l lOcC	hOc4la
26.0	-100.0	2.0	rOcPole	rOcPole	hOc2
-14.0	-98.0	-4.0	lOcPole	lOcPole	hOc1
52.0	-44.0	8.0	rMTG	rMTG/pSTS	-
-56.0	-58.0	36.0	lTPJ	lMTG/pSTS	Area PFm
28.0	-52.0	42.0	rIPS	rSPL	Area hIP1
4.0	-58.0	28.0	rPrc	rPCC	-
52.0	24.0	26.0	rIFS	rIFG	Area45
-46.0	20.0	22.0	lIFG	lIFG	IFS1/IFS2
0.0	20.0	54.0	l pre-SMA	pre-SMA	-
42.0	12.0	30.0	rIFS	rMFG	IFS4
12.0	52.0	16.0	pACC	rMFG	Area p32
8.0	46.0	36.0	r amSFG	rmPFC	-
14.0	28.0	50.0	r pmSFG	rSFG	-
-24.0	24.0	42.0	lMFG	lSFG	-
36.0	2.0	42.0	rMFG	rPrG	-
20.0	-8.0	-14.0	rAm	rAm	Am: SF
-16.0	-6.0	-12.0	lAm	lAm	-

<table>
<tr><td colspan="6" align="center">Reward
Liu et al., 2011</td></tr>
</table>

x	y	z	Macroanatomical Location	Original labeling in the Meta-analysis	Cytoarchitectonic Assignment
12.0	10.0	-6.0	rNAc	rNAc	NAc_fundus
-10.0	8.0	-4.0	lPal	lPal	Striatum_scgp
36.0	20.0	-6.0	raIns	rIns	-
-32.0	20.0	-4.0	laIns	lIns	-

x	y	z	Macroanatomical Location	Original labeling in the Meta-analysis	Cytoarchitectonic Assignment
0.0	24.0	40.0	aMCC	dmPFC	Area 32'
0.0	54.0	-8.0	mOFC	mOFC	Fp2
24.0	-2.0	-16.0	rAm	rAm	Am: LB
4.0	-14.0	8.0	rTh	rTh	Th: Temp
0.0	8.0	48.0	l pre-SMA	SMA	-
8.0	-18.0	-10.0	rBrainstem	rBrainstem	-
2.0	44.0	20.0	rpACC	rACC	Area p32
-24.0	2.0	52.0	lpMFG	lMFG	-
-38.0	-4.0	6.0	lpIns	lIns	Area Id3
24.0	40.0	-14.0	r SOrbG	r midOFC	Area Fo3
-16.0	42.0	-14.0	lSOrbG	l midOFC	-
40.0	32.0	32.0	rpMFG	rMFG	-
-28.0	-56.0	48.0	lIPS	lIPL	hIP3
28.0	-58.0	50.0	rIPS	rAG	hIP3
0.0	-32.0	32.0	PCC	PCC	
-36.0	50.0	10.0	laMFG	lFP	-
-46.0	42.0	-4.0	lIFG	l lOFC	-
30.0	4.0	50.0	raMFG	rMFG	-
-22.0	30.0	48.0	lSFG	lSFG	-

Pain					
Kogler et al., 2015					
x	**y**	**z**	**Macroanatomical Location**	**Original labeling in the Meta-analysis**	**Cytoarchitectonic Assignment**
38.0	18.0	0.0	rIns	rIns	-
52.0	12.0	-4.0	rSTG	rSTG	Area 44
60.0	6.0	2.0	rIFG	rTP	Area 44
22.0	0.0	-4.0	rPal	rPal	-
-38.0	14.0	4.0	laIns	lIns	OP7
-58.0	0.0	6.0	lfOP	lOP4	OP6
-20.0	6.0	2.0	lPut	lPut	Striatum_PM
4.0	6.0	46.0	rSMA	rSMA	Area 24dv
0.0	14.0	36.0	laMCC	lMCC	Areas 24c'v,24c'd
-42.0	-18.0	18.0	lpOP	lOP3	OP3
-54.0	-24.0	24.0	lSMG	lSMG	Area PFop
-36.0	-20.0	2.0	lpIns	lIns	OP7, OP6
-14.0	-12.0	10.0	lTh	lTh	Th: Pref
10.0	-18.0	4.0	rTh	rTh	Th: Pref

x	y	z	Macroanatomical location	Original labeling in the Meta-analysis	Cytoarchitectonic Assignment
56.0	-24.0	24.0	rSMG	rSMG	Area PFop
44.0	-14.0	16.0	r pOP	rOP3	OP3
38.0	50.0	12.0	rMFG	rMFG	-
-24.0	-66.0	-26.0	lCb	lCb	LobuleVI

<div align="center">

Emotion perception

Sabatinelli et al., 2012

</div>

x	y	z	Macroanatomical location	Original labeling in the Meta-analysis	Cytoarchitectonic Assignment
4.0	47.0	7.0	pACC	medPFC	pv24c; pd24cv; pd24cd
42.0	25.0	3.0	rIFG	rIFG	
-42.0	25.0	3.0	lIFG(p.Tr)	lIFG	-
48.0	17.0	29.0	rIFJ	rMFG	IFJ1
-42.0	13.0	27.0	lIFJ	lMFG	IFJ1
-2.0	8.0	59.0	l pmSFG	lSFG	
20.0	-4.0	-15.0	rAm	rAm	Amygdala: SF
-20.0	-6.0	-15.0	lAm	lAm	Amygdala:SF
-20.0	-33.0	-4.0	lHippo	lPHG	.
14.0	-33.0	-7.0	rHippo	rPHG	Subiculum
53.0	-50.0	4.0	rMTG	rMTG	-
38.0	-55.0	-20.0	r aFFG	rFFG	FG3
-40.0	-55.0	-22.0	l aFFG	lFFG	Lobule VI
38.0	-76.0	-16.0	r pFFG	rpFFG	hOc4v
-40.0	-78.0	-21.0	lpFFG	lpFFG	hOc4v
-4.0	52.0	31.0	lamSFG	medPFC	-
36.0	25.0	-3.0	rIns	rOFC	-
-38.0	25.0	-8.0	lIFG(p.Orb)	lOFC	-
2.0	19.0	25.0	aMCC	rACC	Area a24a', a23b'
0.0	-15.0	10.0	lTh	Th	Th: Temporal
-2.0	-31.0	-7.0	Superior Colliculus	Pulvinar	-
-28.0	-70.0	-14.0	lFFG	lFFG	FG1
46.0	-68.0	-4.0	r lOcC	r lOcC	hOc4lp
-48.0	-72.0	-4.0	l lOcC	l lOcC	hOc4lp

<div align="center">

Working Memory

Rottschy et al., 2012

</div>

x	y	z	Macroanatomical	Original	Cytoarchitectonic

				l location	labeling in the Meta-analysis	Assignment
-32.0	22.0	-2.0		l aIns	laIns	-
-48.0	10.0	26.0		lIFG	lIFG (p.Orb)	Area 44
-46.0	26.0	24.0		lIFS	l plPFC	IFS1/IFS2
-38.0	50.0	10.0		lMFG	l alPFC	-
36.0	22.0	-6.0		r aIns	raIns	-
50.0	14.0	24.0		rIFG	rIFG (p.Tr)	Area44
44.0	34.0	32.0		rpMFG	r plPFC	-
38.0	54.0	6.0		raMFG	r alPFC	-
2.0	18.0	48.0		r dmPFC	pmedFC	-
-28.0	0.0	56.0		lSFG	l pSFG	-
30.0	2.0	56.0		rSFG	r pSFG	-
-42.0	-42.0	46.0		lIPS	lIPS	hIP2
-34.0	-52.0	48.0		lSPL	lSPL/IPS	hIP3
-24.0	-66.0	54.0		lSPL	lpSPL	Area7A
42.0	-44.0	44.0		rIPS	rIPS	hIP2
32.0	-58.0	48.0		rIPS	rIPS	hIP3
16.0	-66.0	56.0		rSPL	rpSPL	Area7A
-12.0	-12.0	12.0		lTh	lTh	Th: Pref
-18.0	4.0	6.0		lPutament	lPutamen	Striatum:PoStP
12.0	-10.0	10.0		rTh	rTh	Th: Pref
-34.0	-66.0	-20.0		lFFG/Cb	lCb/FFG	FG2
32.0	-64.0	-18.0		rFFG/Cb	rCb/FFG	FG1
Vigilant Attention Langner et al., 2012						
x	**y**	**z**	**Macroanatomical location**	**Original labeling in the Meta-analysis**	**Cytoarchitectonic Assignment**	
-2.0	8.0	50.0	l pre-SMA	a paracentral lobule	-	
8.0	32.0	46.0	r mSFG	r pmed SFG	-	
0.0	26.0	34.0	l MCC	l/r dorsal MCC	Area 32'	
50.0	8.0	32.0	r IFJ	r IFJ		
40.0	22.0	-4.0	r aIns	r aIns	-	
46.0	36.0	20.0	r MFG	r IFS	-	

-40.0	-12.0	60.0	l PrG	l PrG	-
-46.0	-68.0	-6.0	l IOG	l IOG	hOc4lp; hOc4d; hOc3d
-48.0	8.0	30.0	l IFJ	l IFJ	area 44
62.0	-38.0	17.0	r IPL	r TPJ	area PF
8.0	-12.0	6.0	r Th	r a/mTh	Th: temporal
32.0	-90.0	4.0	r MOG	r MOG	hOc4la
-42.0	12.0	-2.0	l aIns	l aIns	-
-10.0	-14.0	6.0	l Th	l a/m Th	Th: prefrontal
6.0	-58.0	-18.0	r Cb	l/r Cb	lobule V
44.0	-44.0	46.0	r IPS	r IPL	hIP2

Autobiographical memory Spreng et al., 2008					
x	**y**	**z**	**Macroanatomical location**	**Original labeling in the Meta-analysis**	**Cytoarchitectonic Assignment**
-1.0	-53.0	21.0	lPrc	l/rPrc	-
-26.0	-28.0	-17.0	lHippo	lHippo	Subiculum
-49.0	-61.0	31.0	lTPJ	lTPJ	Area PGa
-2.0	51.0	-11.0	lFP	l medPFC	Fp2
-60.0	-9.0	-18.0	lSTS	lSTS/MTG	-
-50.0	27.0	-12.0	lSOrbG	l vlPFC	Fo5
26.0	-33.0	-15.0	rHippo	rpHippo	Subiculum
-1.0	20.0	57.0	lmSFG	MFG	-
55.0	-58.0	30.0	rTPJ	rTPJ	Area PGa
-47.0	9.0	46.0	lPrG	l plPFC	-
-42.0	53.0	7.0	lFP	l lFP	-
26.0	-14.0	-23.0	rHippo	raHippo	DG
52.0	-5.0	-18.0	rMTG	rTP/MTG	-
-39.0	13.0	-41.0	lTP	lTP	-
-38.0	-82.0	38.0	lIPL	lOC	Area PGp
-48.0	29.0	17.0	lIFG	l dlPFC	Area 45
52.0	31.0	-11.0	rSOrbG	r vlPFC	Fo5
-11.0	62.0	9.0	lFP	lmedFP	Fp1
4.0	-8.0	2.0	rTh	rTh	Th: Temporal
-4.0	39.0	16.0	lACC	lrACC	Area pv24c, pd24cv, pd24cd
-5.0	-34.0	36.0	lPCC	lPCC	-

x	y	z	Macroanatomical Location	Original labeling in the Meta-analysis	Cytoarchitectonic Assignment
-29.0	16.0	51.0	lSFG	lSFS	-
31.0	1.0	-26.0	rAm	rAm	Amygdala: LB

<div align="center">

Semantic Memory
Binder et al., 2009

</div>

x	y	z	Macroanatomical Location	Original labeling in the Meta-analysis	Cytoarchitectonic Assignment
-46	-70	21	lIPL	lSTG	Area PGp
-50	-56	31	lAG	lSTG	Area PGa
-64	-44	-4	lMTG	lMTG	-
-47	-24	-17	lMTG	lFFG	-
-55	-3	-24	laMTG	lMTG	-
-7	-57	17	lPrc	lPCC	-
-20	36	44	lSFG	lSFG	-
-31	29	45	lMFG	lMFG	-
-53	26	-1	lIFG	lMFG	Area 45
-39	17	44	lMFG	lIFG	-
53	-59	29	rAG	rSTG	Area PGa
43	-72	31	rpIPL	rMTG	Area PGp
-1	51	-7	medFP	lACC	Area Fp2
-5	56	24	lmSFG	lSFG	Area p32
-31	-34	-16	lFFG	lParaHippo	-
-8	29	-10	sACC	lACC	Area s32
-46	25	23	lIFS	lMFG	IFS1/IFS2
64	-41	-2	rMTG	rMTG	-
-43	-53	55	rIPL	lIPL	Area PFm
-1	-18	40	rMCC	lCC	-
51	20	26	rIFJ	rMFG	IFJ1
64	-38	32	raIPL	rSMG	Area PF
-23	26	-16	rFP	lIFG	Area Fo3

x, y and z coordinates denote the center of gravity in MNI space.

Reference for probabilistic cytoarchitectonic mapping of amygdala and hippocampus (Amunts et al. 2005)); superior parietal cortex (Scheperjans et al. 2008); intraparietal sulcus (Choi et al. 2006); parietal operculum (Eickhoff et al. 2006); ventral extrastriate cortex (Rottschy et al. 2007); dorsal extrastriate cortex (Kujovic et al. 2013); gyrus fusiformis (Caspers et al. 2013); lateral occipital cortex (Malikovic et al. 2016); Broca's regions

(Amunts et al. 1999); Cingulate cortex (Palomero-Gallagher et al. 2015). Cerebellar atlas (Diedrichsen et al. 2009). Thalamic connectivity atlas (Behrens et al. 2003).

Abbreviations: r= right; l= left; a= anterior; p= posterior; s= sub-genual; m/med=medial; Tr.= pars; triangularis; Orb. = pars orbitalis; dmPFC= dorso-medial prefrontal cortex; SMA= supplementary motor area; MCC= middle cingulate cortex; ACC= anterior cingulate cortex; PCC= posterior cingulate cortex; Am= amygdala; Th= thalamus; Hippo= hippocampus; GP/Pal= globus pallidus; Prc= precuneus; mSFG= superior medial gyrus; Nac= nucleus accumbens; Put= putamen; PrG= pre-central gyrus; Ins= insula; IFS= inferior frontal sulcus; IFJ= inferior frontal junction; IFG= inferior frontal gyrus; MFG= middle frontal gyrus; SFG= superior frontal gyrus; OFC= orbito-frontal cortex; SOrbG= superior orbital gyrus; FP= frontal pole; STS= superior temporal gyrus; STG= superior temporal gyrus; MTG= middle temporal gyrus; ITG= inferior temporal gyrus; FFG= fusiform gyrus; SPL= superior parietal lobe; IPL= inferior parietal lobe; IPS= intra-parietal sulcus; fOP= frontal operculum; pOP= parietal operculum; TPJ= temporo-parietal junction; SMG= supramarginal gyrus; AG= angular gyrus; lOcC= lateral occipital cortex; OcPole= occipital pole; MOG= middle occipital gyrus; IOG= inferior occipital gyrus; Cb= cerebellum

Empathy

Static Face Perception

lIFG · lTPJ · lOcC

lMFG · lAm · lOcPole
x=-19

lpSmedG
x=-3

rPrc · rpSmedG · raSmedG · pACC
x=11

rIPS · rOcC · rAm
x=25

rIFG · rMFG · rOcPole
x=45

rMTF · rIFS
x=49

Perception of emotional scenes and faces

lIFJ · lIFG · lIFG (Tri.) · lIFG (Orb.) · laFFG (FG3) · lpFFG (hOc4v)
x= -42

l lOcC · r lOcC · lFFG (FG1) · r aFFG (FG3)
y= -70

lAm · lHippo
x= -20

l pSmedG · l aSmed · lTh · lPulv
x= -2

aMCC · pACC
x= 2

rHippo · rAm
x= 18

r pFFG (hOc4v) · rIns · r aFFG (FG3)
x= 35

rIFJ · rIFG · r lOcC
x= 42

rMTG
x= 53

Reward

lIFG
x=-19

lIPS · laMFG · lpIns · laIns
x=-33

lSFG · lpMFG · lSOrbG
x=-20

aMCC · lSmedG · PCC · mOFC · lPal
x=-5

pACC · rTh · BrainStem
x= 6

Nac
x= 15

rIPS · r aMFG · rAm · rSOrbG
x= 27

rpMFG · rIns
x= 36

Pain

z=20   z=5   z=-2   y=-66

x= 5   x=36

Working Memory

x= -46   x= -30   x= -13   x= 7

y= -69   x= 35   x= 46

Vigilant Attention

x= -45   x= -5   x= 5   x= 37

x=47   x= 64

147

Autobiographic Memory

Semantic Memory
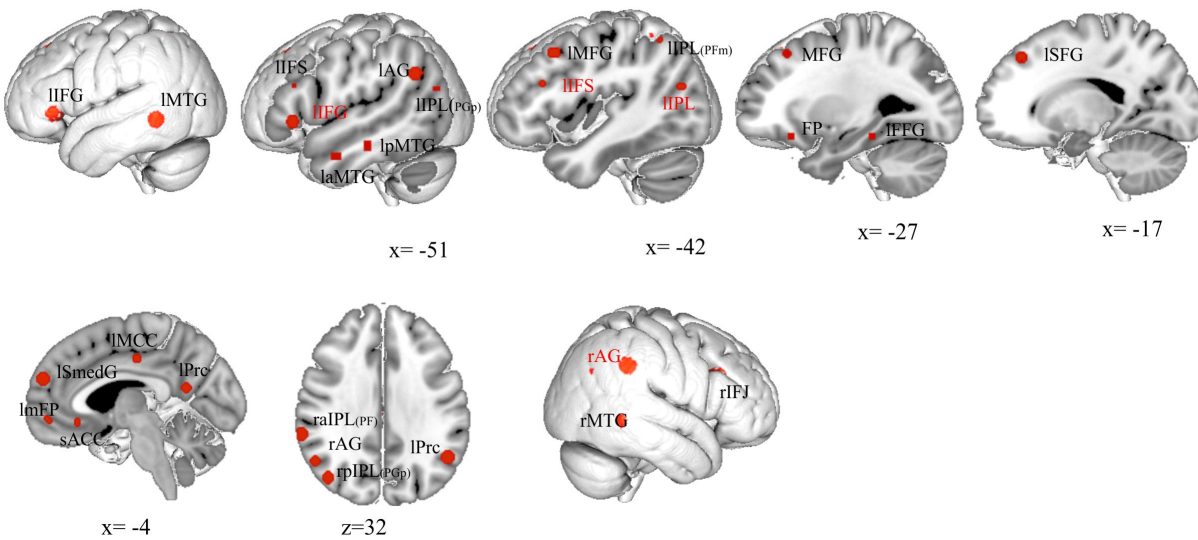
Regions constituting the meta-analytically defined network defined according to the SPM anatomy toolbox 2.1 (Eickhoff et al. 2005, 2007). Red labels indicated regions already defined in previous sections.

Supplement Fig S1: Meta-analytically derived networks
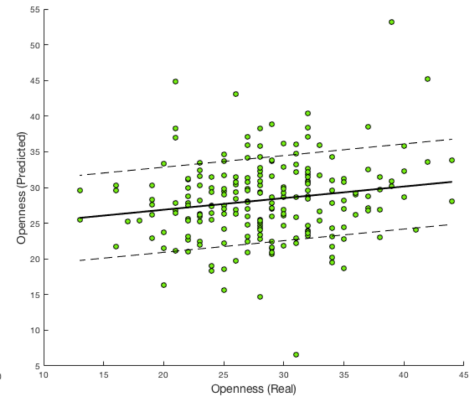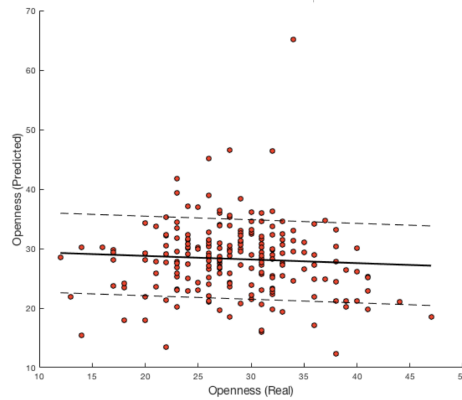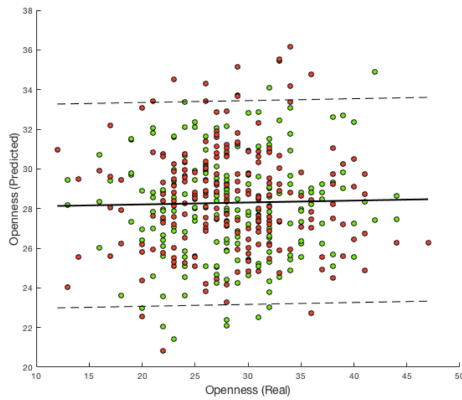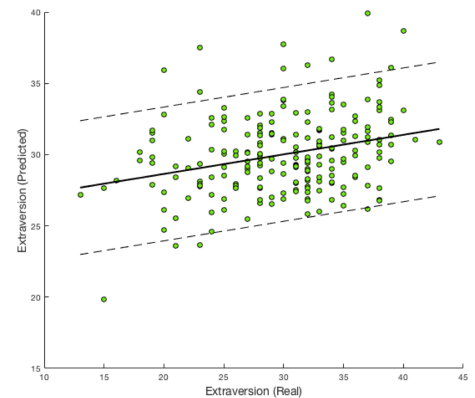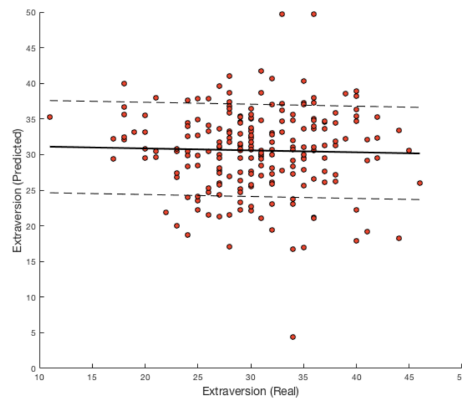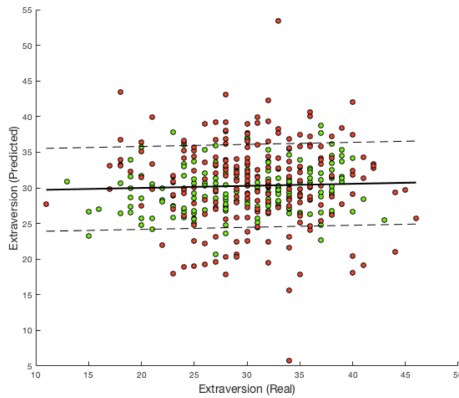
Supplement Fig 2: Comparison of the predictions across groups. Scatter plots of real and predicted personality score in the entire samples (all) as well as for males and females separately. Predictions are reported if they are significant in at least one out of the three groups.

References:

References

Amunts K, Kedo O, Kindler M, et al (2005) Cytoarchitectonic mapping of the human amygdala, hippocampal region and entorhinal cortex: Intersubject variability and probability maps. In: Anatomy and Embryology. pp 343–352

Amunts K, Schleicher A, Bürgel U, et al (1999) Broca's region revisited: Cytoarchitecture and intersubject variability. J Comp Neurol 412:319–341. doi: 10.1002/(SICI)1096-9861(19990920)412:2<319::AID-CNE10>3.0.CO;2-7

Behrens TEJ, Johansen-Berg H, Woolrich MW, et al (2003) Non-invasive mapping of connections between human thalamus and cortex using diffusion imaging. Nat Neurosci 6:750–7. doi: 10.1038/nn1075

Caspers J, Zilles K, Eickhoff SB, et al (2013) Cytoarchitectonical analysis and probabilistic mapping of two extrastriate areas of the human posterior fusiform gyrus. Brain Struct Funct 218:511–526. doi: 10.1007/s00429-012-0411-8

Choi HJ, Zilles K, Mohlberg H, et al (2006) Cytoarchitectonic identification and probabilistic mapping of two distinct areas within the anterior ventral bank of the human intraparietal sulcus. J Comp Neurol 495:53–69. doi: 10.1002/cne.20849

Diedrichsen J, Balsters JH, Flavell J, et al (2009) A probabilistic MR atlas of the human cerebellum. Neuroimage 46:39–46. doi: 10.1016/j.neuroimage.2009.01.045

Eickhoff SB, Paus T, Caspers S, et al (2007) Assignment of functional activations to probabilistic cytoarchitectonic areas revisited. Neuroimage 36:511–21. doi: 10.1016/j.neuroimage.2007.03.060

Eickhoff SB, Schleicher A, Zilles K, Amunts K (2006) The human parietal operculum. I. Cytoarchitectonic mapping of subdivisions. Cereb Cortex 16:254–267. doi: 10.1093/cercor/bhi105

Eickhoff SB, Stephan KE, Mohlberg H, et al (2005) A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. Neuroimage 25:1325–1335. doi: 10.1016/j.neuroimage.2004.12.034

Kujovic M, Zilles K, Malikovic A, et al (2013) Cytoarchitectonic mapping of the human dorsal extrastriate cortex. Brain Struct Funct 218:157–172. doi: 10.1007/s00429-012-0390-9

Malikovic A, Amunts K, Schleicher A, et al (2016) Cytoarchitecture of the human lateral occipital cortex: mapping of two extrastriate areas hOc4la and hOc4lp. Brain Struct Funct 221:1877–1897. doi: 10.1007/s00429-015-1009-8

Palomero-Gallagher N, Eickhoff SB, Hoffstaedter F, et al (2015) Functional organization of human subgenual cortical areas: Relationship between architectonical segregation and connectional heterogeneity. Neuroimage 115:177–90. doi: 10.1016/j.neuroimage.2015.04.053

Rottschy C, Eickhoff SB, Schleicher A, et al (2007) Ventral visual cortex in humans: cytoarchitectonic mapping of two extrastriate areas. Hum Brain Mapp 28:1045–1059. doi: 10.1002/hbm.20348

Scheperjans F, Hermann K, Eickhoff SB, et al (2008) Observer-independent cytoarchitectonic mapping of the human superior parietal cortex. Cereb Cortex 18:846–867. doi: 10.1093/cercor/bhm116

STUDY 4

Evaluation of non-negative matrix factorization of grey matter in age prediction

Deepthi P. Varikuti[1,2,3], Sarah Genon[1,2,3*], Aristeidis Sotiras[4], Holger Schwender[5], Felix Hoffstaedter[1,2,3], Kaustubh R. Patil[1,2], Christiane Jockwitz[1,6,7], Svenja Caspers[1,6,7], Susanne Moebus[8], Katrin Amunts[1,6], Christos Davatzikos[4], and Simon B. Eickhoff [1,2,3*]

[1]*Institute of Neuroscience and Medicine (INM-1, INM-7); Research Centre Jülich, Germany*

[2]*Institute of Systems Neuroscience, Medical Faculty; Heinrich-Heine University Düsseldorf, Germany*

[3]*Institute of Clinical Neuroscience and Medical Psychology; Heinrich-Heine University Düsseldorf, Düsseldorf, Germany*

[4]*Section for Biomedical Image Analysis, Center for Biomedical Image Computing and Analytics, University of Pennsylvania, Philadelphia, USA*

[5]*Mathematical Institute, Heinrich Heine University Düsseldorf, Düsseldorf, Germany*

[6]*C. & O. Vogt Institute for Brain Research, Medical Faculty, Heinrich Heine University, Düsseldorf. Germany*

[7]*JARA-BRAIN, Jülich-Aachen Research Alliance, Jülich, Germany*

[8]*Institute of Medical Informatics, Biometry and Epidemiology, University of Duisburg-Essen, Essen, Germany.*

*Both authors contributed equally to the work

Own contributions   Conception and design of experiment
                           Reviewing and adapting analysis code
                           Statistical data analysis
                           Interpretation of results
                           Preparing figures
                           Writing the paper
                           Total contribution: 80%

Abstract

The relationship between grey matter volume (GMV) patterns and age can be captured by multivariate pattern analysis, allowing prediction of individuals' age based on structural imaging. Raw data, voxel-wise GMV and non-sparse factorization (with Principal Component Analysis, PCA) show good performance but do not promote spatially localized brain components for post-hoc examinations. Here we evaluated a non-negative matrix factorization (NNMF) approach to provide a reduced, but also interpretable representation of GMV data in age prediction frameworks in healthy and clinical populations.

This examination was performed using three datasets: a multi-site cohort of life-span healthy adults, a single site cohort of older adults and clinical samples from the ADNI dataset with healthy subjects, participants with Mild Cognitive Impairment and patients with Alzheimer's disease (AD) subsamples. T1-weighted images were preprocessed with VBM8 standard settings to compute GMV values after normalization, segmentation and modulation for non-linear transformations only. Non-negative matrix factorization was computed on the GM voxel-wise values for a range of granularities (50 to 690 components) and LASSO (Least Absolute Shrinkage and Selection Operator) regression were used for age prediction. First, we compared the performance of our data compression procedure (i.e., NNMF) to various other approaches (i.e., uncompressed VBM data, PCA-based factorization and parcellation-based compression). We then investigated the impact of the granularity on the accuracy of age prediction, as well as the transferability of the factorization and model generalization across datasets. We finally validated our framework by examining age prediction in ADNI samples.

Our results showed that our framework favorably compares with other approaches. They also demonstrated that the NNMF based factorization derived from one dataset could be efficiently applied to compress VBM data of another dataset and that granularities between 300 and 500 components give an optimal representation for age prediction. In addition to the good performance in healthy subjects our framework provided localized brain regions as the features contributing to the prediction, thereby offering further insights into structural changes due to brain aging. Finally, our validation in clinical populations showed that our framework is sensitive to deviance from normal structural variations in pathological aging.

1 Introduction

The structural dynamics of the human brain during adulthood is a highly complex process. Machine-learning algorithms have been used to capture the multivariate pattern of structural brain changes (Franke et al. 2010) that relate to age with a brain-based age prediction framework. By suggesting an age for any individual based on his/her brain's structural scan, such approaches can provide new insights into brain plasticity, into accelerating cerebral aging, as well as into the influence of several variables such as genes, pharmacological intervention and cognitive training in both healthy and clinical populations. Voxel based morphometry (VBM) is one of the most commonly used methods to measure grey matter volume (Good et al. 2001). It provides non-negative measures, which convey biologically meaningful information and capture brain changes related to age and pathology, as well as brain plasticity related to training (Good et al. 2001; Tisserand et al. 2002; May 2011). Previous studies have shown that machine-learning methods applied to VBM data allow prediction of brain age (Franke et al. 2010). In these studies, brain age was estimated by applying a support vector machine approach on the high dimensional voxels' data (Erus et al. 2015). However, in voxel-wise representation of structural data, features may convey redundant information and/or noise and may promote overfitting due to a higher number of features relative to the number of subjects (Guyon and Elisseeff 2003; Hua et al. 2009; Mwangi et al. 2014). To address this issue, Franke et al. (2010) examined brain age prediction based on the simple and fast application of the principal component analysis (PCA) to the data and subsequent brain age prediction with a relevance vector machine approach. This combination allowed them to predict the brain age with an absolute error of 5 years. Ever since, Franke et al. 2010's framework has been employed to investigate other concepts in relation to healthy aging (such as different age groups i.e., children and adolescents (Franke et al. 2012), gender differences (Franke et al. 2014)), as well as differences between healthy aging and various neurocognitive deviancies (such as cognitive impairments (Gaser et al. 2013) and psychiatric disorders (Koutsouleris et al. 2014)).

Most of the above-mentioned studies have implemented principal component analysis (PCA) to counter the curse of dimensionality associated with multivariate analysis of neuroimaging data (Franke et al. 2010; Franke et al. 2012; Franke et al. 2013; Liem et al. 2017). PCA decomposes the entire non-negative representation into a low rank approximation with a combination of positive and negative weights (Jolliffe 2002), which does not promote spatially localized components. Furthermore, the signed components within the PCA decomposition engage complex cancellations during the reconstruction of the original representation. Therefore, the use of PCA-based dimensionality reduction on brain voxels hardly results in interpretable components, which can in turn prevent the interpretation of a predictive model based on PCA-derived components. Non-negative matrix factorization (NNMF) is an alternative decomposition method promoting spatially localized representation

that has gained more attention in the past years. NNMF can factorize a given dataset into low-ranking approximations capturing a parts-based representation (Lee and Seung 1999). The non-negativity constraint leads to only additive combinations of the components, which allows the factorization to reconstruct the original high dimensional data from the parts-based representation. As a result, NNMF provides a more interpretable factorization compared to standard decomposition approaches such as PCA and ICA (Independent Component Analysis) (Lee and Seung 1999; Sotiras et al. 2015). Recently, Sotiras et al. (2015) investigated the application of NNMF to neuroimaging data, by decomposing the structural MRI data with an extended version of NNMF, the orthonormal projective non-negative matrix factorization (OPNMF). This approach provided components that could be considered as a biologically more meaningful parts-based representation of the brain as compared to more standard approaches such as PCA and ICA. Accordingly, OPNMF promotes spatially localized brain components for post-hoc examinations. Hence, OPNMF could open new perspectives for dimensionality reduction of VBM data, in particular in a prediction framework. However, to the best of our knowledge, these perspectives have remained unexplored. Therefore, the current study aimed at examining the application of OPNMF to VBM data in a brain-age prediction framework. To note, we have used the term NNMF when denoting to the whole family of the technique, whereas, OPNMF when referring to the more the specific variant, which we have employed in this study.

We first compared the performance of OPNMF-based factorization to the performance of plain VBM data for age prediction. Then, in order to provide direct comparison with previous studies, we evaluated different strategies combining either PCA or OPNMF as a data compression approach with either LASSO or RVM as sparse regression models. In addition to the sparseness inducting methods described above, several parcellations of the human brain have been proposed in the last two decades (Eickhoff et al. 2017), which could potentially offer another efficient approach for data compression into localized spatial units for age prediction. In particular, many whole-brain parcellations have been derived from voxels/vertex functional signal at rest (RS, e.g.: Bellec et al. 2010; Craddock et al. 2012; Gordon et al. 2016; Schaefer et al. 2017) and such RS-based parcellation has been used for the compression of RS functional connectivity (RSFC) data in a brain age prediction framework (Liem et al. 2017). However, we assume that such a representation based on functional parcellation is, by nature, less optimal than a representation based on the structural properties of the voxels as used in the current VBM-based framework. To investigate this hypothesis, we compared the pattern of representation, as well as the prediction performance of our data reduction approach OPNMF capitalizing on structural covariance with an independent brain representation derived from resting-state functional data in healthy adults.

OPNMF is computationally more expensive than popular decomposition methods such as PCA (see methods). Nevertheless, transferring the factorization derived from one dataset

onto another dataset could save this computational cost. Furthermore, using factorization from an independent dataset for training or testing a prediction model can assess the robustness of the model. We, therefore, evaluated the transferability of the OPNMF onto an unseen dataset, that is, we examined the transferability of the components derived from one dataset onto an independent (new) dataset, hence avoiding the time-consuming step of factorization in the new dataset. Importantly, transferring the already computed components onto a new dataset is particularly useful in clinical and research practices, as the datasets often come from different sites and scanners and may have different demographic characteristics. Recently, Liem et al. (2017) suggested that combining datasets from different protocols could reduce the bias of the predictive model towards the characteristics of a single protocol. Therefore, the effect of data acquisition and demographic heterogeneity on the transferability of the components is an important aspect to evaluate in the perspective of application of our framework in future studies. Here, we examined a dataset from a uniform protocol constituting older subjects (age range 55-76) vs. a heterogeneous multi-site dataset whose age range covers the adult life span (19-81, Fig. 1A)). Thus, we assessed the performance of the prediction model trained on a dataset compressed using its own factorization, as well as, when this dataset was compressed based on an independent factorization (that is, when the dataset has been projected onto a factorization derived from a different dataset).

In addition, the difference in the sample characteristics of the two cohorts further offer the opportunity to investigate the extrapolation of the prediction model trained on one dataset onto an independent dataset. That is, in the present study, we investigated both, the transferability of the components among datasets and the generalization of the prediction among datasets, on the age prediction performance. Relatedly, one crucial objective in age prediction is the identification of aging trajectories deviating from normal range, i.e., pathological aging. Previous studies have shown dramatic brain structural alterations in patients with pathological aging such as Mild Cognitive Impairment (MCI) and Alzheimer's Disease (AD) resulting in systematic overestimation of their age by an algorithm trained on healthy populations (Davatzikos et al. 2009; Gaser et al. 2013; Moradi et al. 2015). Therefore, as a validation of our framework for clinical research, we further evaluated its performance in age prediction of healthy and clinical samples from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database.

To sum up, in this study, we aimed to evaluate a new framework for brain-age prediction, which used dimensionality reduction of VBM data using OPNMF followed by a sparse regression model. In order to evaluate the advantages and limitations of this framework over the other approaches proposed in the previous studies, we compared the performance of our model with 1) model based on voxel-wise VBM data (uncompressed VBM data), 2) model based on PCA data reduction and 3) model based on data reduction based on RS-based

parcellation. In the sake of reducing computational cost in future studies, we examined the transferability of the OPNMF between two independent datasets differing in demographic characteristic and acquisition protocols. Importantly, the localized properties of the components in our framework allowed us to explore brain regions contributing to the predictiveness in the healthy samples. Finally, we tested the performance of our prediction model on a clinical sample, in order to validate the predictive utility of our framework in clinical research.

2 Material and Methods

2.1 Sample characteristics and preprocessing

We used structural MRI data from two large, independent datasets. The first was obtained from the population-based 1000BRAINS study (Caspers et al. 2014) and represents a single-site assessment of 693 older adults (age: 55-75 years; 53% males) using the same imaging protocol for all subjects. The other "MIXED" dataset consists of 1,084 healthy adults (age: 18-81 years; 51% males) that were derived by pooling data from many different individual studies at various sites (Fig. 1A; for further details see Supplementary methods). Furthermore, in order to validate our framework of age prediction on clinical data, we included a dataset from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (www.loni.ucla.edu/ADNI). This ADNI dataset sample included 244 cognitively normal elderly subjects (HC, age: 55-90; 48% males), 64 mild cognitively impaired (MCI) subjects (age: 55-87; 60% males), and 163 Alzheimer's disease (AD) subjects (age: 56-91; 56% males), for further details see Supplementary methods.
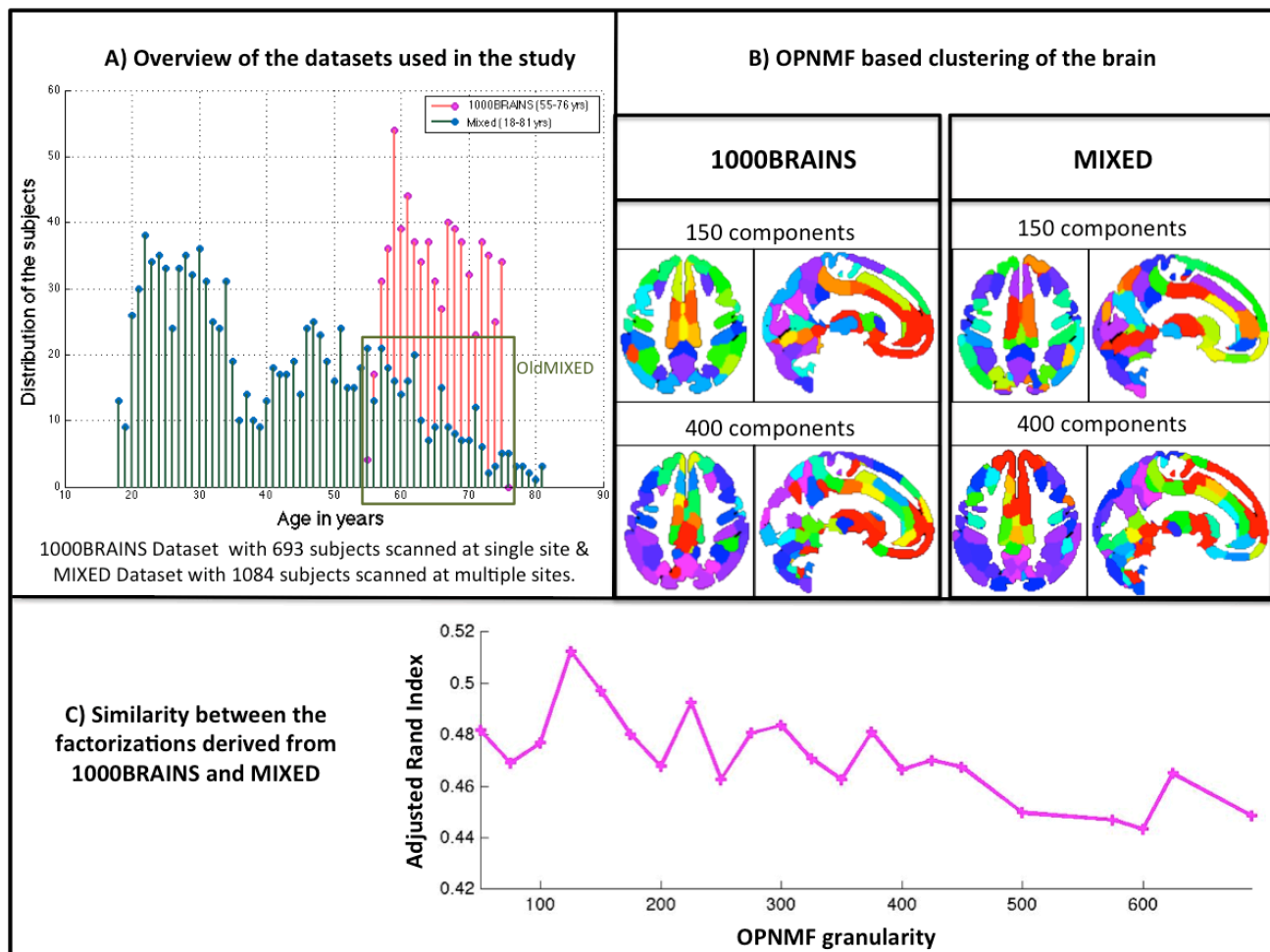

Fig 1. The two healthy datasets and the non-negative matrix components derived the these datasets. A: Overview of the sample characteristics of the two datasets (i.e., range of age

distributed in each dataset, as well as the scanner protocol). B: Brain spatial representation of the factorization derived from the two datasets at two different resolutions. C: Similarity between the factorizations derived from the two datasets.

Structural MRI data was preprocessed with the VBM8 toolbox (http://www.neuro.uni-jena.de/vbm8) to derive voxel-wise grey matter volumes for each subject of the two datasets using standard settings. T1-weighted structural brain images were normalized by the high-dimensional DARTEL normalization (Ashburner 2007) combined with tissue class segmentation and bias field correction. The normalized grey matter segments were modulated for non-linear transformations only and smoothed with an 8-mm FWHM Gaussian kernel. The local grey matter volumes (following adjustment of head size given that the affine part of the registration did not enter the modulation) were then extracted in a whole-brain grey matter mask (with a threshold of 0.2 to eliminate the voxels with partial volume effect (Ashburner et al. 1985)) and for each sample individually stored in a Number of subjects by Number of voxels matrix (with Number of voxels = 344,383). These matrices provided the input for the age-prediction model based on the full (uncompressed) VBM data as detailed in section 2.4 and the input to which matrix factorization (i.e., non-negative matrix factorization and principal component analysis) and resting-state (RS) based parcellation (see below) were applied.

2.2 Data reduction
2.2.1 Orthonormal projective Non-negative matrix factorization
We used the same orthonormal projective non-negative matrix factorization (OPNMF) approach (Yang et al., 2007; Yang and Oja 2010) as described by Sotiras et al. 2015. OPNMF factorizes a data matrix 'X' into two non-negative sub matrices (W and H) representing the sparse components (the dictionary i.e., W of size, Number of voxels by Number of components) and the subject-specific loading coefficients (H of size, Number of components by Number of subjects) in the ensuing low-rank space, $\min_{W \geq 0, H \geq 0} \parallel X - WH \parallel_F^2$, which minimizes the squared Frobenius norm (i.e., reducing the reconstruction error), subject to the conditions $H = W^T X$ and $W^T W = I$ where, $\parallel . \parallel_F^2$ referred to the squared Frobenius norm and I denotes the identity matrix.

To summarize the factorization process, W is first initialized through non-negative double singular value decomposition (NNSVD; cf. Boutsidis and Gallopoulos 2008)). Later, W is iteratively updated with the multiplicative update rule, until it converges to an optimum solution. The multiplicative update rule is modified as reported by Yang and Oja (2010), to satisfy the additional constraints of an orthonormal projection basis function, $W'_{ij} = W_{ij} \frac{(xx^T W)_{ij}}{(ww^T xx^T W)_{ij}}$, where, $i = 1 \ldots$Number of voxels, $j = 1 \ldots$Number of components. Finally, projecting X onto W to obtain a solution that minimizes the reconstruction error yields H.

Following OPNMF, the VBM data are represented by two matrices, denoting the sparse components (W) and the corresponding subject-specific loading coefficients (H). The former (W) represent the latent structure in the data, the latter (H) represents the individual volumetric data in the low-rank space spanned by these components and provides the features for the age-prediction model. Of note, the highest possible OPNMF granularity is the lowest dimension of the input matrix (X) (which in our case is the number of the subjects in 1000BRAINS dataset being the smallest sample size). Accordingly, in this study, to explore the effects of different granularity, i.e., number of components, on prediction accuracy, we computed and evaluated compressions employing 50 to 690 components in steps of 25.

2.2.2 Principal component analysis (PCA)

PCA is one of the most commonly used dimensionality reduction techniques and, accordingly, has been used in previous studies examining age prediction based on structural MRI data. In order to provide a direct comparison of the OPNMF's performance with the previous investigations, we ran PCA on the voxel-wise VBM data by using the PCA function implemented in MATLAB 2014. This transformed the high-dimensional voxel wise data (i.e., X) into low-rank approximations using an orthogonal linear transformation. The resulting PCA based low-rank approximations represented the principal components of the data (computed by solving an eigenvalue problem) arranged in descending order of the variance explained by each component (Jolliffe 2002). The subject-specific loading coefficients were obtained by projecting the high-dimensional voxel wise data onto the component space (eigenvectors, i.e., PCA based low-rank approximations) thereby providing the features for age-prediction model. Finally, we computed and evaluated the effect of PCA compressions on prediction accuracy, in the range of granularity aforementioned for OPNMF (i.e., 50 to 690 components in steps of 25.)

2.2.3 Resting-state (RS) based parcellation

Recently, Schaefer et al. 2017 reported a parcellation based on RS fMRI providing neurobiologically-valid brain parcels by capitalizing on a new hybrid approach integrating the local gradient approach for boundary-mapping with a global similarity approach. As this atlas does not cover subcortical and cerebellar structures, we added these from another widely used RS fMRI parcellation (BASC, Bellec et al. 2010). This resulted in a whole brain parcellation of 470 parcels that was used here as an alternative dimensionality reduction approach for VBM data. For each subject, an average grey matter volume within each parcel was computed and used as inputs for the age-prediction model.

2.3 Sparse regression model

We primarily used LASSO (Least Absolute Shrinkage and Selection Operator) for learning a (sparse) linear regression model predicting the subjects' age from their structural data as

compressed in the loading coefficients (as implemented in the 'glmnet' package, https://www.jstatsoft.org/article/view/v033i01 (Tibshirani 1996)). LASSO regulates the parameters (alpha and lambda) to optimize the sparsity and the complexity of the regression model to improve the performance (i.e., prediction accuracy) and interpretability of the model (Zou and Hastie 2005; Zhang and Huang 2008). An inner loop was incorporated to optimize the hyper-parameter (lambda). LASSO with alpha set to 0.99 and lambda that gives minimum mean cross validation error of the inner loop was employed for predicting the age in our study.

As an alternative approach to LASSO, Relevance Vector Machine (RVM; Tipping 2001) has been commonly implemented by the previous studies exploring prediction of age using structural MRI. Therefore, we in this study performed an additional comparison between LASSO and RVM regression models. For doing so, statistical learning of the sparse regression model employing RVM was implemented using the SparseBayes package (http://www.miketipping.com/index.htm). This approach uses a probabilistic Bayesian framework with specific priors over the parameters, which favors sparse prediction model. The algorithm iteratively and automatically optimizes the parameters and hyper parameters, hence reducing prior control on the parameters. As kernel, we chose a multivariate zero-centered Gaussian with standard deviation estimated by the algorithm. This RVM implementation from the SparseBayes package has been shown to improve the initialization procedure, which maximizes the likelihood function and hence accelerates the procedure (Tipping and Faul 2003).

2.4 Prediction analyses

Previous studies of age prediction from MRI data in life span cohorts have used linear regression model (Franke et al. 2010; Franke et al. 2012; Gaser et al. 2013; Mwangi et al. 2013; Franke et al. 2014; Koutsouleris et al. 2014; Erus et al. 2015; Liem et al. 2017). For the sake of comparability, we likewise used a (sparse) linear regression model for predicting the subjects' age from their structural data as compressed in the loading coefficients. Furthermore, in the present study, combining sparse decomposition method with a sparse regression model came with the advantage of providing an anatomically well interpretable model for estimating age based on a limited number of spatially compact structural features.

Model generalization was evaluated by 10-fold cross-validation. That is, the dataset was randomly split in ten equal parts that each, in turn, served as the test set for the model fitted on the remaining $9/10^{th}$ of the data. To reduce dependency on the cross-validation split, this procedure was replicated 100 times. Splitting the dataset into ten equal parts has been initiated every time within each repetition, which allowed us to train the model on different training samples in each repetition. Prediction accuracy was quantified by the mean absolute deviation (across subjects) between real age and predicted age (averaged across repetitions),

and also, the correlation between the real age and that average of the predicted (across repetitions) in previously unseen subjects from their VBM data.

2.5 Assessed prediction approaches

We note that performing OPNMF only on the training dataset in each cross-validation fold would be computationally expensive and hence practically infeasible. But prediction performed on loading coefficients obtained from the OPNMF decompositions including the entire sample (including the $1/10^{th}$ that is denoted the test-set in the respective fold) could bias the cross-validation towards overly optimistic performance estimates. Given this consideration would also hold for all future uses of our approach, we were particularly interested in investigating whether components derived from one dataset were also good encoders (representative) for the structural features of another dataset.

Consequently, we performed cross-validation analyses using the loading coefficients derived from OPNMF of that particular dataset as (an overly optimistic) reference (Fig. 2: A&B) but importantly focused on assessing the possibility to predict subjects' age after projecting the raw VBM data on the component space estimated for the other dataset (Fig. 2 C&D). That is, we derived the OPNMF components of the 1000BRAINS dataset, and performed cross-validation within the MIXED dataset projected onto the components estimated from the 1000BRAINS. This approach has the advantage that the subjects in the test set were truly independent and have not been involved in any prior processing steps. In addition, we could investigate the effects of dataset (in-) homogeneity, as the 1000BRAINS data comes from a single-site study with uniform protocol, whereas the MIXED dataset was deliberately chosen to be very heterogeneous. We would thus expect that the components derived for the MIXED dataset show a better generalization than those from the 1000BRAINS dataset, i.e., projecting the 1000BRAINS data onto the components from the MIXED dataset will yield prediction models that are closer in performance to the (optimistically biased) analysis with the projection of the data components derived from the MIXED data itself.

Later, we also tested whether the actual models transfer between datasets by deriving the OPNMF components in one dataset (e.g., MIXED), fitting the sparse regression model in that same dataset (MIXED), projecting the other dataset (1000BRAINS) onto the factorization of the former (MIXED) and applying the prediction model trained on that (MIXED) data (Fig. 2: E, F&G). In this context, we note that the 1000BRAINS dataset has a more restricted age-range (55-76) than the MIXED dataset (18-81). Therefore, we evaluated the model transfer between the portion of MIXED subjects corresponding to 1000BRAINS age range (55-75; i.e., OldMIXED), see Fig. 2: E&F. The computational times for each prediction approach at different levels of granularity are reported in supplementary material (Table S3).
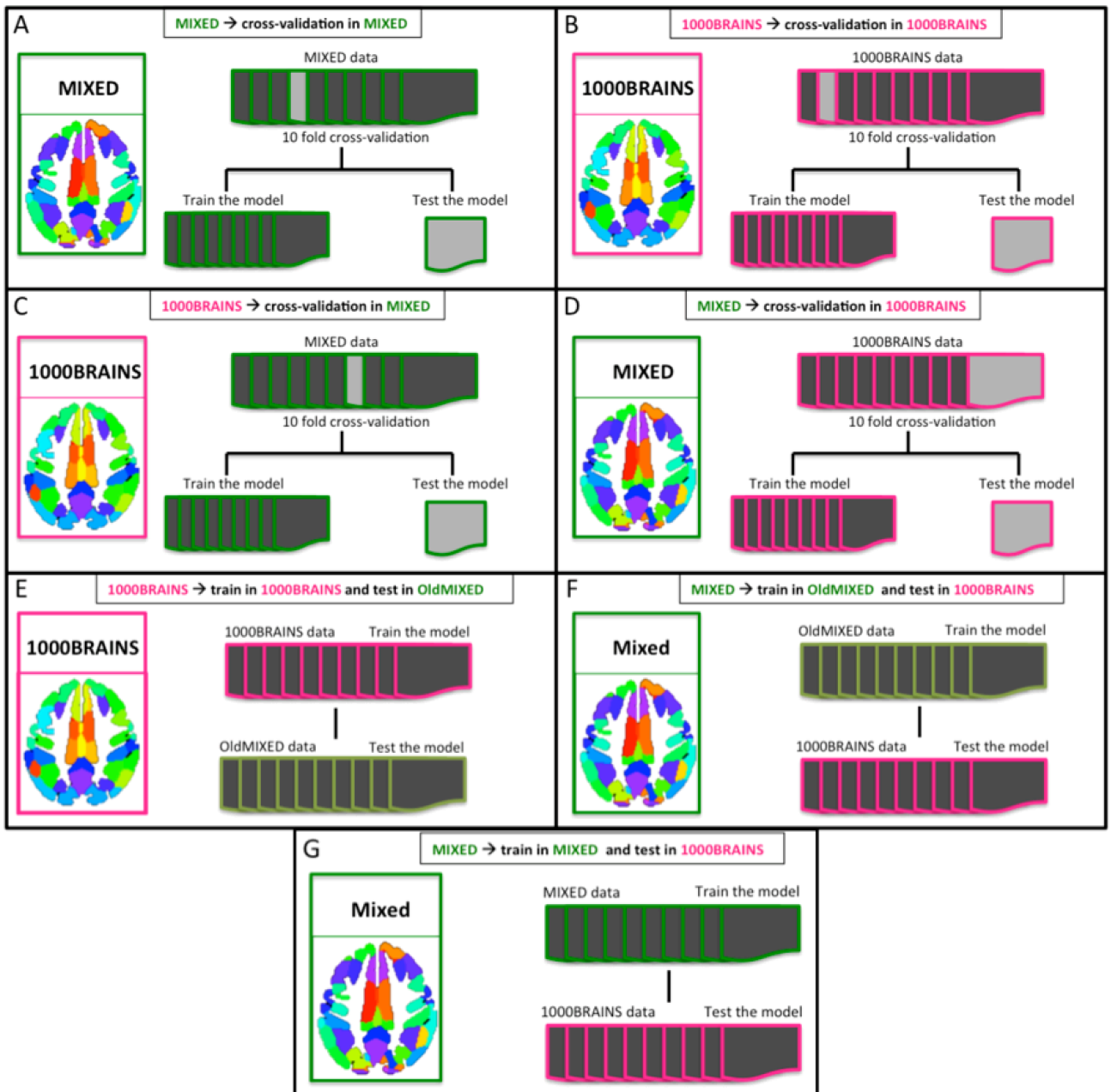
Fig 2. Different prediction approaches evaluated in the study. A&B illustrate the procedure of a 10 fold cross-validation when compressing a dataset using the components derived from itself (A) MIXED dataset (B) 1000BRAINS. C&D illustrate the procedure of a 10 fold cross-validation performed on features extracted by using the components derived from the other dataset (C) MIXED dataset projected on to 1000BRAINS based factorization, and (D) 1000BRAINS projected on to factorization derived from MIXED. E,F&G illustrate the approaches utilizing an independent dataset to validate the model trained on the dataset compressed using the components derived from itself (E) training the model on 1000BRAINS dataset projected on to its own factorization and later validate the model on OldMIXED dataset, (F) training the model on OldMIXED dataset projected on to OldMIXED based factorization and later validate the model on 1000BRAINS and (G)

training the model on MIXED dataset projected on to MIXED based factorization and later validate the model on 1000BRAINS.

Finally, our age estimation framework was validated using the ADNI database. Here we compared the estimated BrainAGE between healthy controls (HC=, subjects with mild cognitive impairment (MCI) and Alzheimer's disease (AD), given that apparent older brains have been previously demonstrated in the latter two groups. As AD patients sample mainly consisted of older subjects, the prediction model was trained on the aforementioned samples of elderly subjects (i.e., 1000BRAINS or MIXED_55-90). In detail, all data (training sample and ADNI) were projected onto the factorization derived from the respective training sample (either the 1000BRAINS or MIXED). The model was trained on the each of the training sample (1000BRAINS or MIXED_55-90 (i.e., Subjects above 55 years from MIXED dataset)) and then evaluated it in the ADNI data. In line with previous studies (Davatzikos et al. 2008; Franke et al. 2010; Franke and Gaser 2012; Moradi et al. 2015), we hypothesized, that for the ADNI controls, the brain age gap estimation (BrainAGE), i.e., the difference between the predicted age and the chronological age, should be centered around zero. In turn, MCI subjects and in particular AD patients were expected to show an increased BrainAGE score.

2.6 Identification of the regions influencing the prediction

As noted above, combining a sparse decomposition yielded compact regional modes with a sparse regression model (LASSO) has the advantage of providing regionally specific relevant features. As a final step allowing the neurobiological interpretation of our age-prediction model, we identified those parts of the brain that underpinned the reported predictions. As we performed 100 replications of a 10-fold cross-validation, in total 1000 models were computed per granularity (number of components). We then quantified the contribution of each component by the fraction of these 1000 models in which the coefficients assigned by the predictive model for the respective component was non-zero. The components that contributed in at least 95% of all models were identified as the components that were robustly part of the predictive model (Fig 3 (5)). Concretely, we first identified the components consistently contributing to the prediction as defined by non-zero beta value in 95% of the models. Second, the components were mapped to the brain space at each respective level of granularity. That is, we built a "contributor map" at each level of granularity, in which the voxel values represent their (binary) contribution (Fig 3 (5)). Combining those maps (by summing up the values) resulted in a contributor "summary" map in which a non-zero value represents a contribution in at least 95% of the 1000 prediction models) and higher value represent more overlap across different granularities (Fig 3 (6)). As we found that prediction performance stabilizes after around a granularity of 300 components (Fig 6), only the contributor maps of granularity > 300 components were merged into a summary map. Given that the relationship between the OPNMF components at high

granularities could be hierarchically inconsistent, this approach yielded a higher effective resolution of the relevant brain areas than the actual granularity of the factorizations itself, and hereby alleviated the reliance of the spatial inference on any particular set of components.
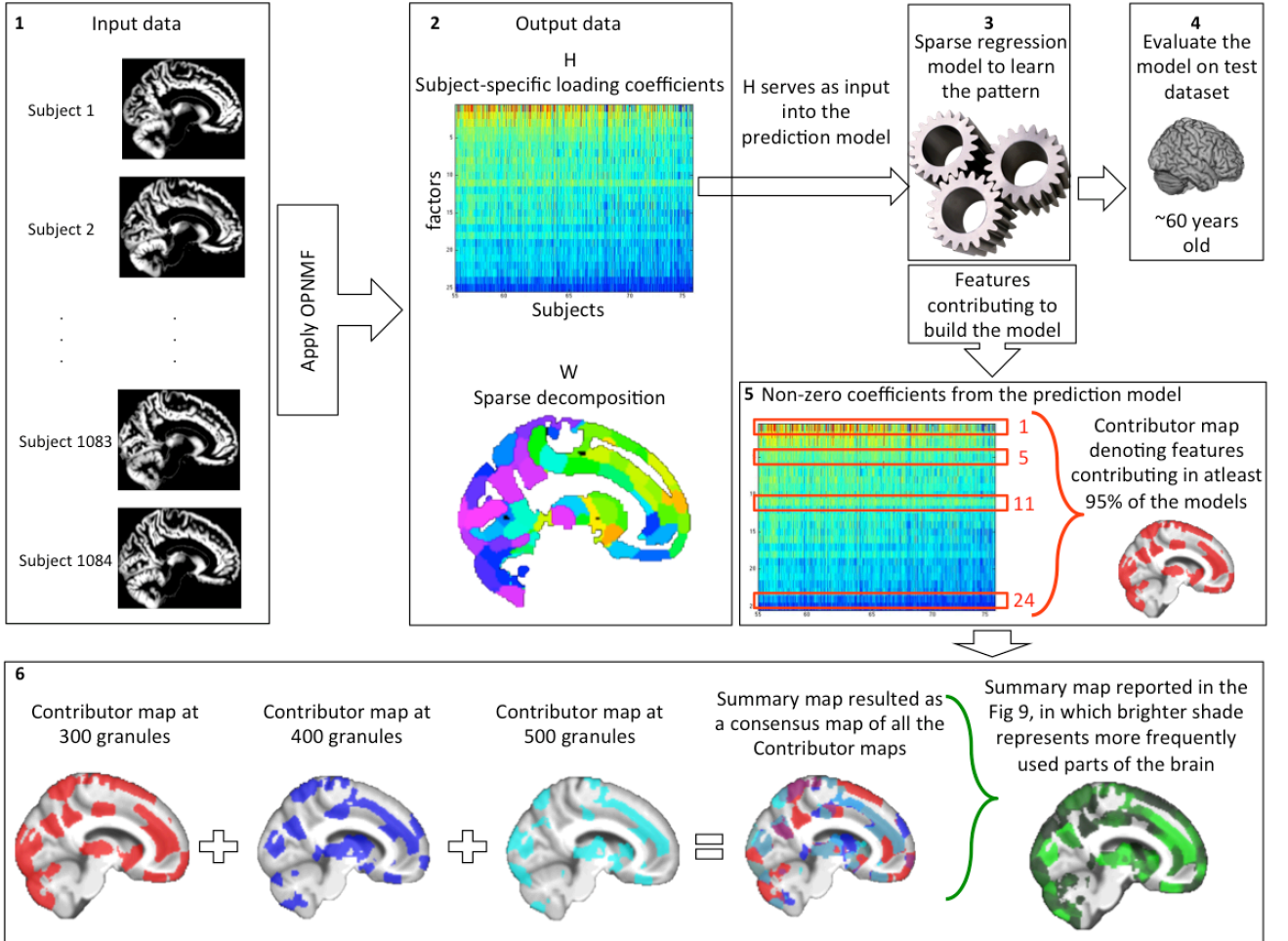

Fig 3. Main processing steps for age prediction based on GMV and the post-hoc examination of regions contributing to the prediction. 1) Voxel-based morphometric (VBM) data for each subject are used as input for OPNMF 2) Following OPNMF, the VBM data are represented by two matrices, denoting the corresponding subject-specific loading coefficients (H) and the sparse components (W). 3) Application of sparse regression model, in which H provides the features for the prediction model 4) Evaluation of the prediction model using a test sample (different prediction models described in section 2.5 & Fig 2). 5&6) Identification of the regions contributing in the prediction analysis; 5) First the respective components with non-zero coefficients assigned by the prediction models were identified. Then, we built a "contributor map" at each level of granularity, in which the voxel values represent their (binary) contribution in at least 95% of the models. 6) Combining those maps (by summing up the values) resulted in a contributor "summary" map in which a non-zero value represents a contribution in at least 95% of all the prediction models) and higher value represent higher overlap across different granularities. As our analyses revealed that prediction performance

stabilizes around 300 components (Fig 6), only the contributor maps of granularity > 300 components were merged into the summary map.

3 Results

3.1 Brain age estimation using the uncompressed VBM data

Training LASSO models on the full, i.e., uncompressed voxel-wise VBM data allowed to predict the age of previously unseen subjects with relatively high accuracy. For the 1000BRAINS data, the mean absolute error (MAE) between real and predicted age of the test set was 3.4 years. While for the MIXED dataset, the MAE was 4.9 years. While these numbers compare favorably with previous reports, Fig. 4 illustrated the critical drawback of using sparse regression models on voxel-wise data. That is, isolated voxels scattered across the brain were selected as relevant features by the prediction model. In addition to being computationally prohibitive, the ensuing models are basically uninterpretable as the predictions were driven by individual voxels (Fig. 4).
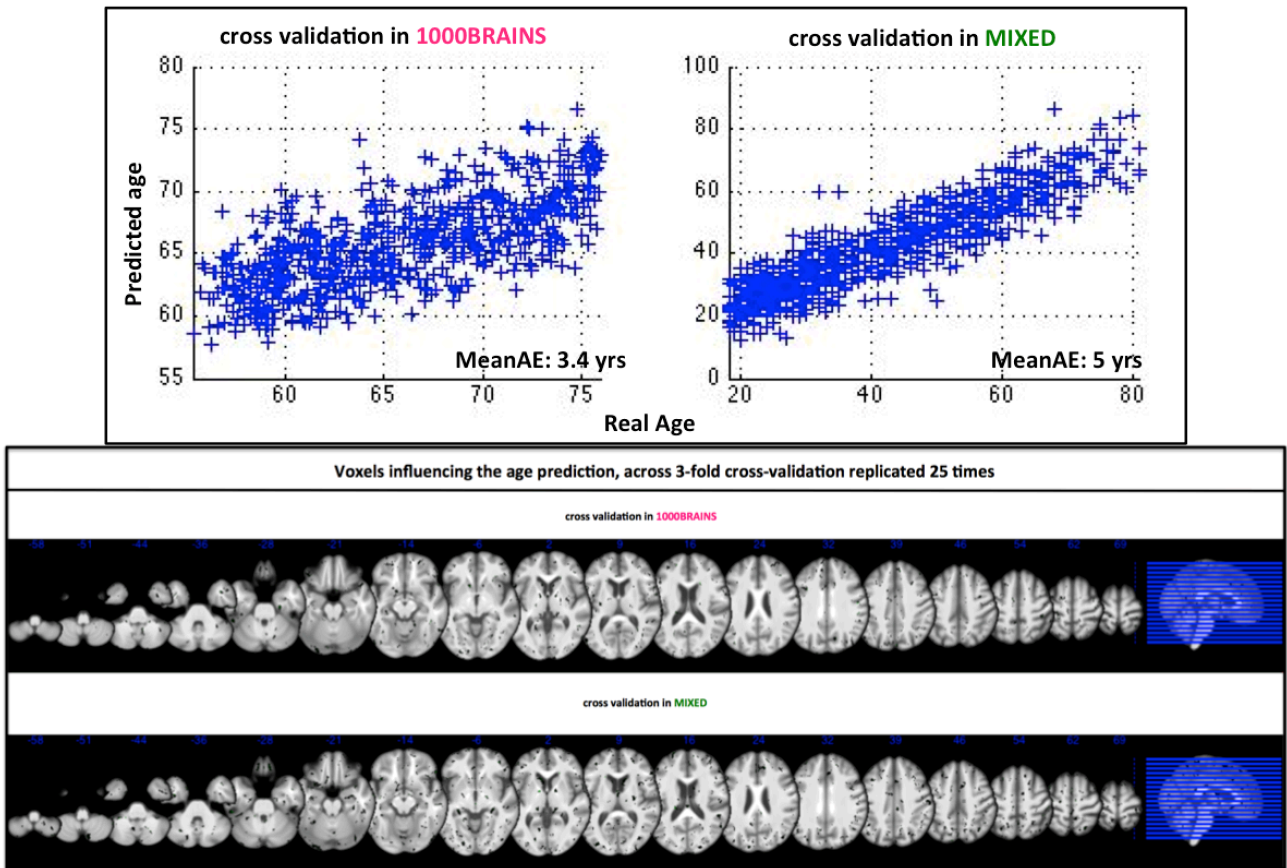

Fig 4. Chronological age plotted against the age predicted using the high-dimensional VBM data. The lower figure exhibits the isolated voxels that contributed in the prediction analysis. Here, the voxels, which contributed in all the models across 25 replications of 3-fold cross-validation, are displayed.

3.2 Compression of Brain age estimation using different compression methods and sparse regression models

Fig 5 illustrated the performance of each of the four combinations of approaches. Across different cross-validation approaches (Fig 5A), OPNMF either slightly outperformed or remained analogous to PCA, especially at higher level of granularity. In particular, when the factorization has been transferred from one dataset to another dataset (Fig 5A: plots on the right compared with plots on the left), OPNMF reported more accurate age prediction with stable performance across different levels of granularity compared to PCA. Thus, we could infer from our results that OPNMF derived from one dataset could provide a better representation of the structural data of an independent dataset than PCA. With respect to the sparse regression approach, LASSO and RVM resulted in comparable cross-validation accuracies, but LASSO was shown to yield superior performance when predicting age across samples (Fig 5B), irrespective of the employed factorization. Additionally, the application of LASSO together with OPNMF performed better than the other combinations in most of the scenarios, supporting the combination of LASSO with OPNMF for age prediction analyses. Accordingly, we focused on investigating the brain age prediction using LASSO sparse regression model, in the subsequent analyses (such as, comparison of OPNMF with a previous RS-parcellation, examination of the OPNMF transferability, identification of the localized features contributing to the prediction analyses in healthy datasets and finally validation of our framework (i.e., combination of OPNMF with sparse regression model) in a clinical dataset).

As suggested by our comparative analyses, based on the prediction accuracies reported by the LASSO (sparse) regression model, we compared the performance of the OPNMF factorization with a RS-parcellation of the brain (Bellec et al. 2010; Schaefer et al. 2017; cf Section 2.2.3). As illustrated in Fig 6, at comparative levels of granularity (i.e., 475 OPNMF factors vs. 470 brain parcels), the age prediction model tended to be more accurate when the data have been compressed with OPNMF than when the data have been compressed based on an independent representation derived from RS fMRI signal. Nevertheless, it has to be noted that the latter also compressed data into localized brain parcels, which by itself showed surprisingly good performance, suggesting that different spatial representations into local components can be efficient (see discussion). In the scope of the current study, altogether, our preliminary comparative analyses supported the use of OPNMF for data compression in an age prediction framework.
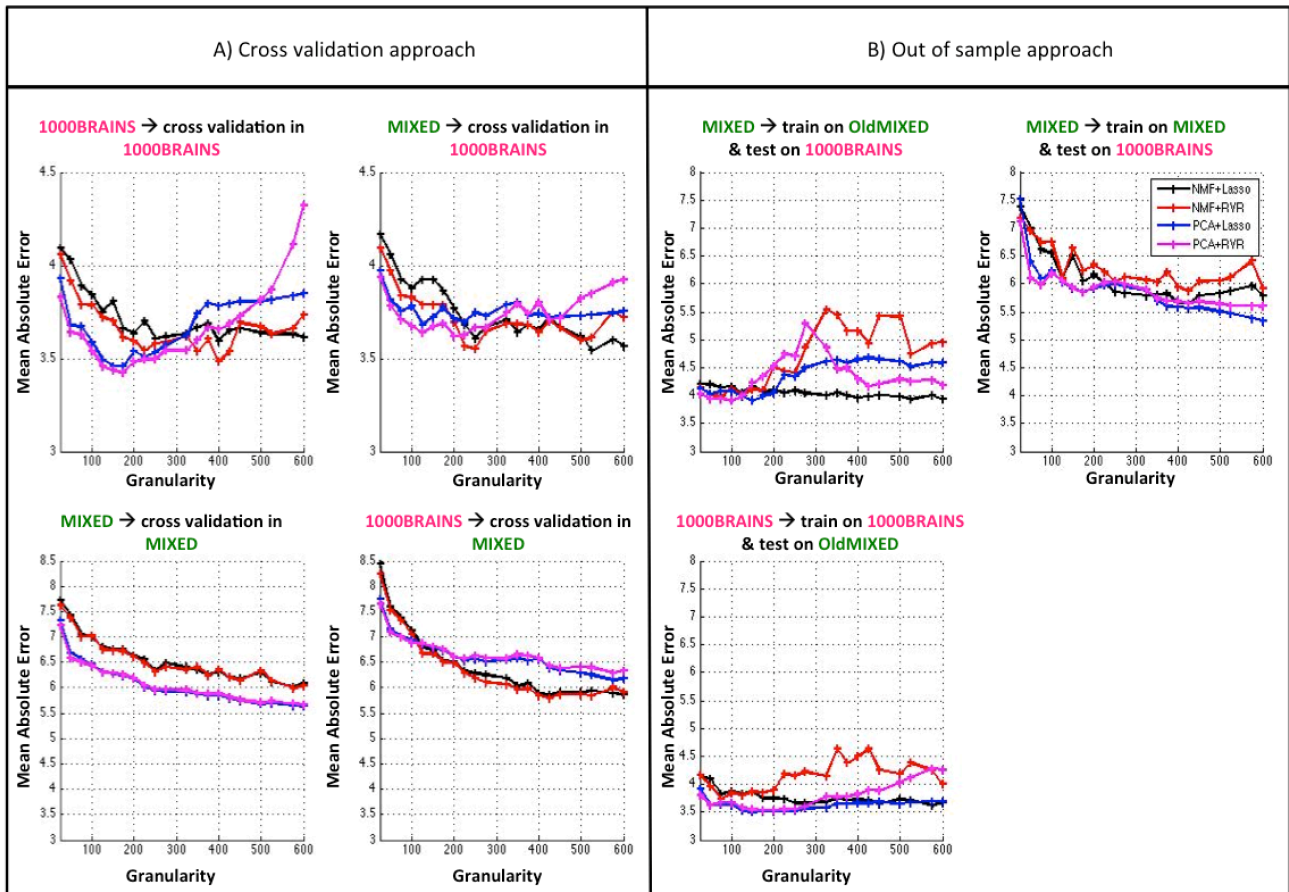
Fig 5. Illustration of the mean absolute error of the prediction models using different compression methods with various sparse regression models (LASSO+OPNMF in black; LASSO+PCA in blue; RVM+OPNMF in red; RVM+PCA in magenta) at different levels of granularity, and separately for each prediction approach. A) represents the approach where 10 fold cross-validation is performed and B) illustrates the approaches in which an independent dataset is used to validate the model trained on the dataset compressed based on the components derived from itself.

3.3 Influence of different datasets on the OPNMF & age prediction

As previously reported, OPNMF provided sparse and spatially compact components, which essentially reflect local structural covariance (Fig. 1B). While not the primary focus of this work, we noted that labeling each grey-matter voxel by the most strongly reflected component, provided a map of the human brain that in many aspects seemed to resemble those derived from other modalities. Across the whole range of granularity, although there seems to be a decent agreement, these maps were slightly different between both investigated datasets (1000BRAINS and MIXED) as reflected in the adjusted rand index (aRI, Fig. 1C). This latter quantifies the similarity between the clusters (Hubert and Arabie 1985; Santos and Embrechts 2009)) between the respective parcellation and can range between +1 and -1, with 1 reflecting perfect spatial correspondence, 0 indicating spatial agreement with certain probability, and smaller than 0 representing disagreement which is worse than contingency (Hubert and Arabie 1985)). However and more importantly, both factorization

(1000BRAINS and MIXED) at the similar level of granularity (i.e., 475 granules) showed good convergence with the RS-parcellation (470 parcels) with, respectively, aRI = 0.28 and aRI = 0.29 (Fig. 6).

3.4 Brain age estimation using the OPNMF-compressed VBM data

Considering the models based on the loading coefficients for components derived from the (full) data of the same sample (rendering data compression not independent from the latter cross-validation), several important observations can be made. First, while very low-rank approximations only yielded poor prediction accuracy, the mean absolute error (MAE) quickly declines with increasing granularity, i.e., higher number of components. Once the number of components passes approximately 300 – 400 (Fig. 6), however, prediction accuracy seems to remain stable or at best improve asymptotically.
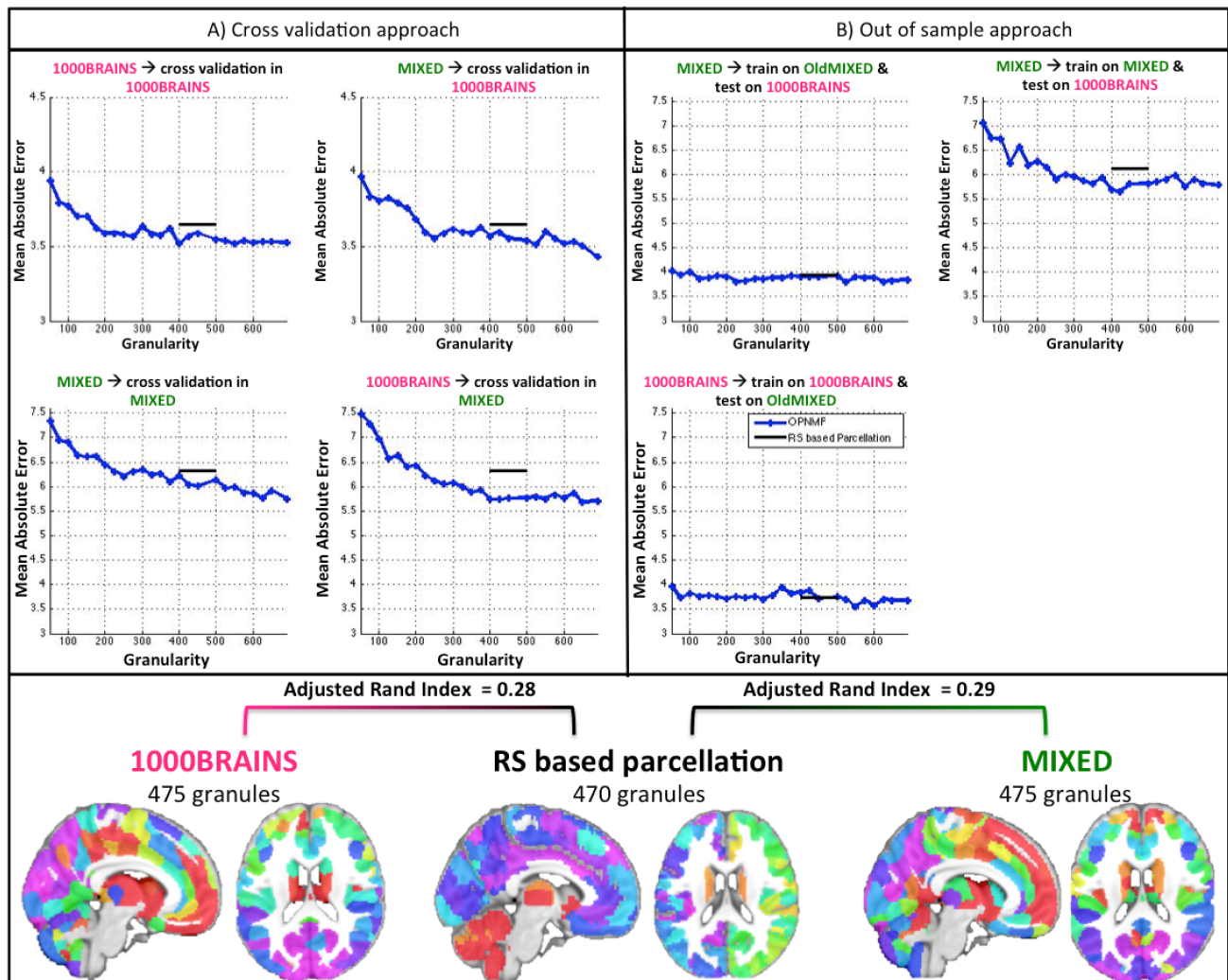

Fig 6. Illustration of the mean absolute error of the prediction models using different spatially localized compression models (OPNMF and RS-parcellation) with sparse (LASSO) regression model at different levels of granularity, and separately for each prediction approach. A) represents the approach where 10 fold cross-validation is performed and B)

illustrates the approaches in which an independent dataset is used to validate the model trained on the dataset compressed based on components derived from itself.

3.4.1 Model validation within the same dataset using a cross-validation approach

In details, these unbiased models yielded an overall MAE of 3.6 years (males: 3.7 and females: 3.6) and an overall correlation of 0.65 (male: 0.62 and female: 0.61) between real and predicted age in the 1000BRAINS data (using components derived from the MIXED dataset; Table S1 & Table S2). For the MIXED data, we found an overall MAE of 6.1 years and a correlation of 0.88 (MAE of 6 and $r = 0.88$ in the males and MAE of 6.3 and $r = 0.86$ in the females) when using components derived from the 1000BRAINS dataset (see Table S1 & Table S2 for detailed results). Further examining the prediction performance across the different scanning sites forming the MIXED dataset (16 different sites) revealed that in most of them (14 sites) the MAE varied between 5 to 7 years and the MAE from the two other sites were 4 years and 9 years (of note, the scanning protocols used in these latter remained analogous to the 14 other sites, i.e., we did not note any specific technical factor accounting for the differences in prediction accuracies). Overall, these results showed the stability of our prediction framework across genders and scanning sites.

3.4.2 Model validation with prediction in independent datasets

Transfer of the whole pipeline (factorization and model training) was evaluated by predicting the age of the subjects in the respective other, independent sample (Fig 6B & Fig 7B). In our study, transferability of the prediction model was evaluated in two different aspects, extrapolation of the prediction model onto an independent dataset, which differs in subjects' demographic characteristics, such as age, and then onto an independent dataset, which differs in scanner protocols. In the context of dataset (in-) homogeneity from different age groups, models trained on broader age range of the heterogeneous dataset (MIXED) showed reduced precision of age prediction in an independent dataset (1000BRAINS) while the model trained on a restricted age range for this restricted heterogeneous dataset (OldMIXED) was more accurate in predicting the age of the latter independent dataset (1000BRAINS). In the context of dataset (in-) homogeneity from different protocols, our results surprisingly showed that models trained on single-site study (1000BRAINS) also performed efficiently, when predicting the age of highly heterogeneous dataset (OldMIXED). In contrast, models trained on the 1000BRAINS data consisting exclusively of older subjects showed a very poor performance when trying to predict age of the younger subjects in the MIXED sample (Fig S2). While the model correctly predicted the young subjects to be younger than the young examples in the training set, it was grossly inaccurate in predicting how much younger they actually were. Put pointedly, having no information about how a 20-year old brain looks like, a model trained on subjects aged between 55 and 76 can only derive that the subject in question should be younger than the youngest it has seen in the training data (Fig S2). Thus, testing for generalization of the model to an independent dataset showed good prediction

accuracy for the subjects within the training sample's age range (Fig 6B & Fig 7B), but also indicated that the prediction model cannot extrapolate to the subjects whose age is (far) beyond the training samples age distribution (Fig S2).

As ultimately one main application of our framework will be research in clinical populations, we also tested our framework in the ADNI dataset. Here, the mean BrainAGE scores (reflecting, for each subject, the discrepancy between brain-based estimated age and chronological age) was zero in the healthy control group (for models trained on either the 1000BRAINS or the MIXED_55-90 datasets). In contrast, BrainAGE scores were 6.2 years (for models trained on 1000BRAINS) and 5.4 years (for model trained on MIXED_55-90) in the MCI group, indicating that these subjects' brains looked about 5-6 years older. Finally, the BrainAGE scores reached 8.5 years (for models trained on 1000BRAINS) and 10.7 years (for models trained on MIXED_55-90) in the group of patients diagnosed with Alzheimer's disease. These results, illustrated in Fig 8 demonstrated that our framework can accurately capture the range of normal structural variation relating to age in healthy subjects and building on this normal range, captures dramatic deviance in both patients with Mild Cognitive Impairment and patients with Alzheimer's disease.
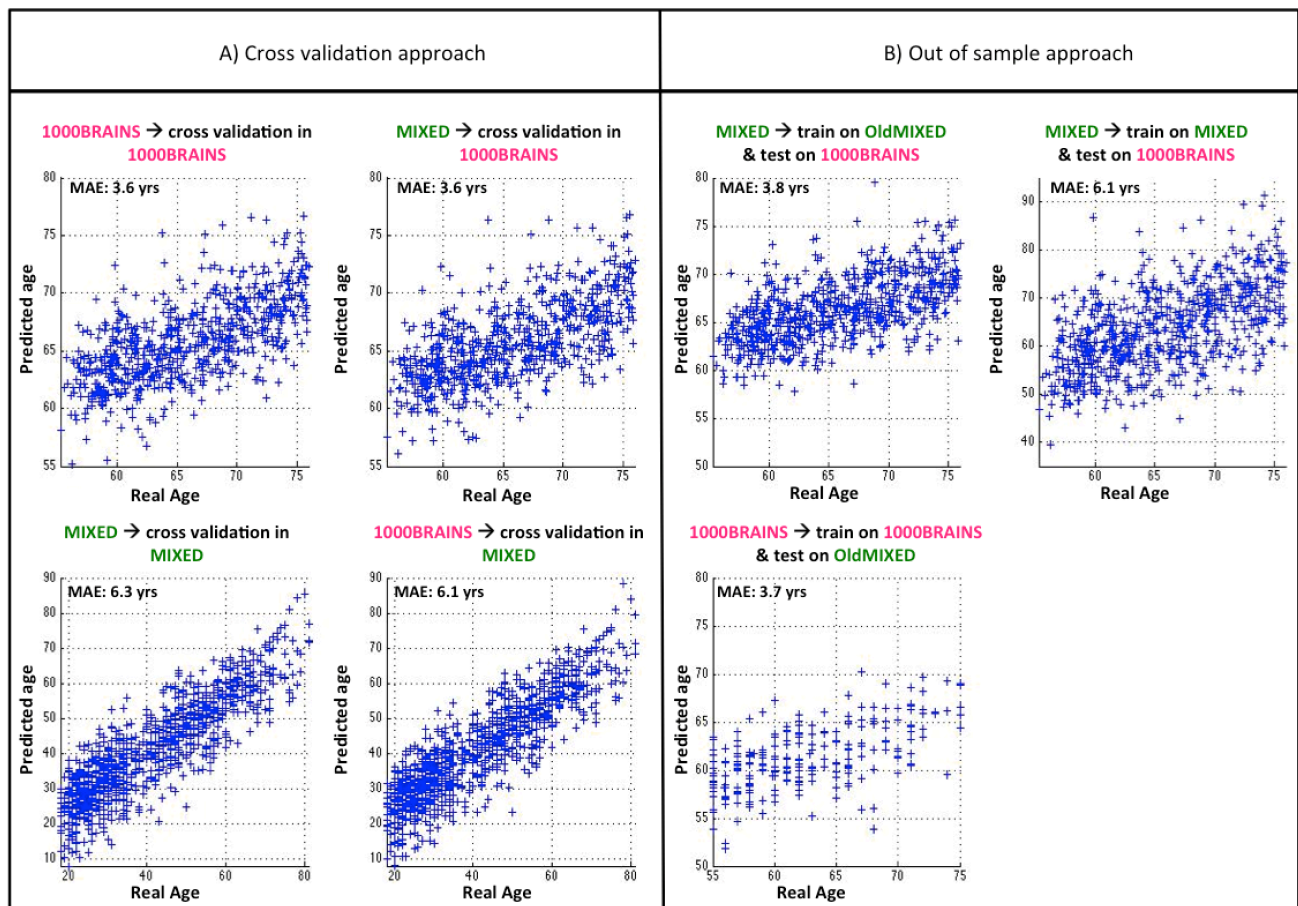

Fig 7. Chronological age plotted against the age predicted using the VBM data compressed

with OPNMF. The predicted age plotted in this figure is an average of the predicted age across different levels of granularity.
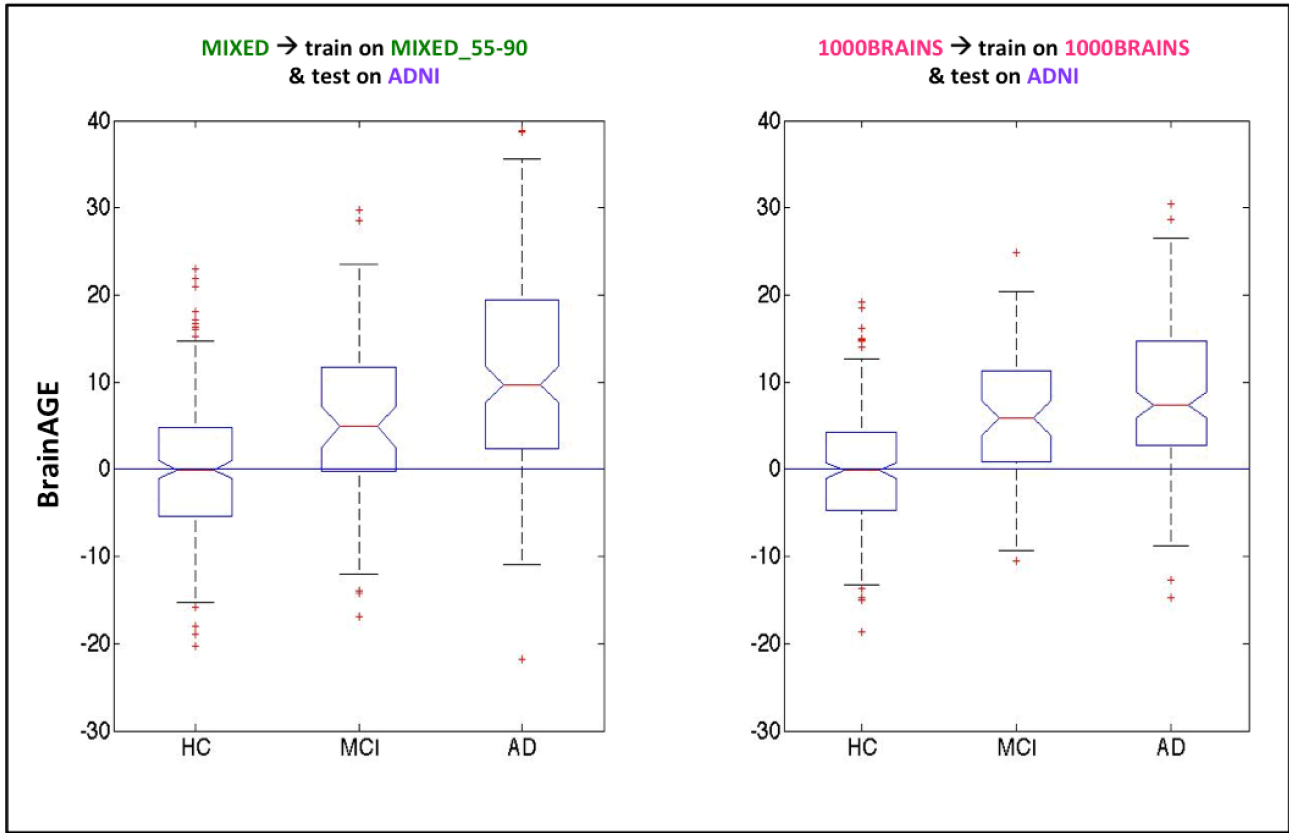

Fig 8. Validation approach in ADNI samples. BrainAGE scores (reflecting the difference between predicted age and the chronological age) are showed for all the three subsamples (i.e., Healthy controls (HC), Mild Cognitive Impairment (MCI), Alzheimer's Disease (AD). The left plot refers to the approach in which the model was trained on OldMIXED sample compressed using factorization derived from the whole MIXED sample. The right plot refers to the approach in which the model was trained on 1000BRAINS compressed using factorization derived from 1000BRAINS.

3.5 Identification of the regions influencing the prediction

The framework we examined in this paper, i.e., applying a sparse regression model onto the sparse decompositions, should yield rather confined and hence neurobiologically interpretable maps of brain regions contributing to the age prediction. In more detail, as previously noted, the OPNMF components themselves were circumscribed rather than representing a mixture of voxel-wise positive and negative weights as would be the case for PCA (cf. Sotiras et al., (2015)). LASSO then selected a small number of these spatially confined components for the actual prediction. This allowed us to identify, which brain regions consistently contributed to the age estimation. As illustrated in Fig 9, more brain

regions were engaged in estimating age in the MIXED as compared to the 1000BRAINS dataset, which could be expected given the much broader age distribution.

More specifically, the regions contributing to the prediction model in the 1000BRAINS cohort (older subjects) included regions around the central sulcus, the inferior temporal cortex, the occipital and posterior temporal cortices and area 44. Regions contributing to the predictions in this older adult cohort also included bilateral midline areas such as, the superior medial frontal gyrus, the medial fronto-orbital regions, the anterior and middle cingulate cortices and the retrosplenial cortex. Furthermore, the pattern of regions weighting in the prediction model in this cohort further included bilateral subcortical regions such as the thalamus, the basal ganglia and the posterior hippocampus, as well as the bilateral cerebellum. On the lateral surface, the pattern included regions in prefrontal regions (frontal areas anterior to the precentral gyrus), oribitofrontal regions and temporal poles. In contrast, the pattern of regions for age prediction in the MIXED dataset (heterogeneous dataset covering the whole life span) was less spatially specific, covering most of the brain lateral surface bilaterally (including for example the whole bilateral middle and superior frontal gyri, as well as the bilateral posterior superior and inferior parietal cortices), almost the entire medial structures, and, the bilateral anterior hippocampus and amygdala. In other words, the prediction models of age in this heterogeneous dataset built on most of the brain regions.
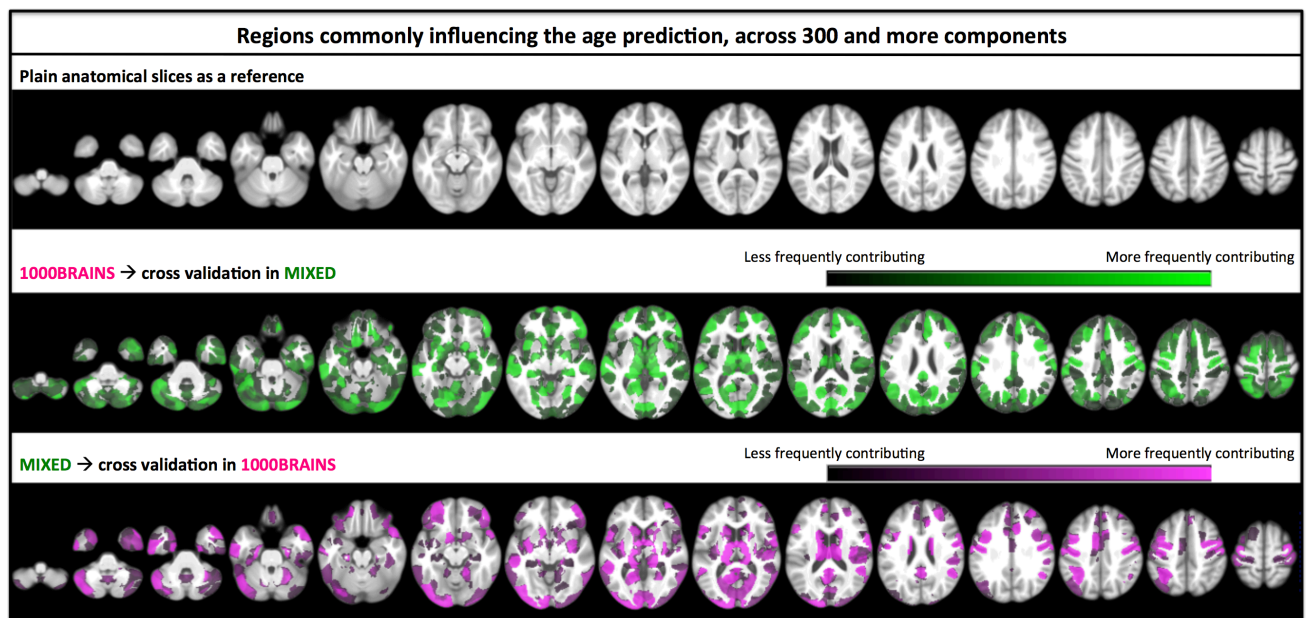

Fig 9. Summary map of the regions that contributed in the prediction analysis when performing 10-fold cross-validation and compressing the dataset using the components derived from the other dataset, in which brighter shade represents more frequently used parts of the brain. Plain anatomical slices are displayed as reference in the top raw. The middle raw illustrates the MIXED dataset compressed with the 1000BRAINS-based factorization

while the bottom row illustrates 1000BRAINS dataset compressed with the MIXED-based factorization.

3.6 Supplementary Analysis

Spatial smoothing on the VBM data promotes homogeneity of the data by attenuating small differences between individuals. In turn, age prediction may rely on those subtle effects. Thus, we also evaluated, whether the subjects' age could be predicted based on the unsmoothed VBM. As could be expected from the aggregation of individual voxels into components, refraining from smoothing prior to projection resulted in highly similar results as shown above for the smoothed data (see Supplement, Table S4 & Fig. S1).

4 Discussions

In this study, we showed that non-negative sparse coding through the combination of data compression using OPNMF with LASSO regression could predict age of previously unseen subjects in an unbiased manner from structural neuroimaging data. Several key observations emerged from this work. i) The precision of age prediction compares well to that based on uncompressed, i.e., voxel-wise VBM data and to that based on non-sparse factorization (PCA). ii) Even though the components estimated for the two datasets differed from each other, the (unbiased) prediction accuracy after projection onto the respective other set of components is only slightly worse than the (biased) accuracy obtained when performing factorization of the entire dataset that was later used for cross-validation. iii) OPNMF-based brain partitions show some convergence with an independent parcellation based on resting-state (RS) fMRI, but the former gave slightly better prediction performance iv) Finally, in contrast to approaches used in previous age prediction studies, combination of data compression using OPNMF with sparse (LASSO) regression yields a superior interpretability of the weight maps allowing interpretations about the mechanisms underlying the prediction.

4.1 Prediction from uncompressed VBM data

Our results showed that age prediction of unseen subjects using the full (uncompressed) VBM data reported only slightly better prediction accuracies than one based on the (OPNMF) compressed (Table S1). This comparable level of performance for compressed and uncompressed data has also been observed in previous brain age studies employing PCA compression (Franke et al. 2010; Liem et al. 2017). However, in the current study, predicting age using sparse regularization (LASSO) prediction model on uncompressed VBM data is highly inefficient in terms of memory usage, especially for large datasets (for example, MIXED dataset with 1,084 subjects). In particular, the memory load of this high-dimensional approach (>700 subjects x 344383 voxels) only allowed a 3-fold cross-validation on a high-performance server. While high dimensional voxel wise data could also lead to overfitting of prediction model, due to the larger number of features than subjects (i.e., several models potentially could fit the same data), comprehensively investigating this issue was not possible in the present study due to the computational limitations. Beside this still open issue, the recent availability of MRI data in very large sample sizes, i.e., for thousands of subjects (e.g., (Miller et al. 2016)) and the growing interest for the prediction of phenotype or behavioral measures from MRI data, dramatically underpin the need of dimensionality reduction preserving prediction accuracy (Davatzikos 2016).

The key limitation of voxel-wise analysis, however, is the poor interpretability of the relevant features. As shown in Fig. 4, the sparse regression model on the voxel-wise data in our study highlighted isolated voxels scattered over the brain as relevant features for predicting subjects' age. Nevertheless, the individual anatomical correspondence of a

particular voxel chosen by the prediction model, can be variable across subjects (Davatzikos 2004). In addition, LASSO regression is known to perform reliable feature selection, providing that the features have followed "irrepresentable condition" (Zhao and Yu 2006). That is, features should be independent of each other in order to obtain reliable outcomes. Therefore, when LASSO is applied to voxel-wise VBM data, the isolated voxels from the highly correlated voxel-wise VBM data contributing to the prediction cannot really be interpreted. In other words, voxel-wise sparse regression models pose a decoding problem (Kampa et al. 2014). Thus, the poor interpretability of prediction models based on raw VBM data (Lakkaraju et al. 2016), in addition to their computational costs, advocate for data compression, ideally with a factorization approach that offers interpretability of the representations such as the current implementation of OPNMF on VBM data for prediction of brain age.

4.2 Compression of Brain age estimation using different compression methods with various sparse regression models

When comparing the performance of OPNMF with PCA, particularly at higher level of granularity, our results demonstrated that OPNMF either slightly outperformed or remained analogous to PCA. Any data reduction procedure aims to address the curse of dimensionality without any loss of information. In this context, both PCA and OPNMF provide low rank approximations representing the most influential structure within the original data, however, each decomposition method captures different aspects of the similar information (PCA captures the components with the most variance explained across the dataset, while OPNMF captures the spatially localized components that consistently co-vary across the dataset), leading in the present study to comparable performance of both approaches in age prediction. Importantly, our results also further showed that OPNMF provided more stable performance at high granularities (> 200), when compression is transferred across datasets (Fig 5A: cf. right vs. left plots). This finding confirms previous hypotheses that the 'projectivity' of OPNMF supports the efficient transferability of the factorization onto a new unseen dataset (Yuan et al. 2007). Therefore, we would argue that OPNMF, compared to PCA, enhances the generalizability of the low rank approximations onto an independent dataset. Thus, OPNMF not only promotes localized brain representation, but also yields the more generalizable low-rank approximation than PCA.

Our evaluation further revealed that LASSO regularization performed either similarly or slightly better than RVM. Both LASSO and RVM yield sparse regression models with the advantage of performing feature selection by capitalizing only on the features that improve the prediction accuracy and allow comparable accuracies. However, an additional argument for the use of LASSO, this model allows the selection of the regularization parameter. Hence, LASSO optimizes the trade-off between stability and interpretability of the prediction model (i.e., optimizing the sparsity) by tuning the regularization parameter (i.e., alpha),

which linearly combines the L1 and L2 penalties (cf. Zou and Hastie 2005 for more technical details). Therefore, the LASSO regression model can convert the sparse regression model into a purely non-sparse model (using the elastic net regularization model) and can therefore be considered as a relatively more flexible regression model than RVM. Furthermore, Bunea et al. 2011 demonstrated that LASSO could be implemented in many conditions including when the feature size is exceeding the sample size for prediction analyses. Indirect support for this property of LASSO can be seen in top left plot of Fig 5B, in which LASSO works particularly better than RVM after crossing the granularity level of 250 (with the training sample size approximating 230 subjects in this specific case). As reduction techniques have shown best prediction accuracies at higher level of granularity and given previous considerations, we focused on LASSO for subsequent prediction analyses.

4.3 Compression of OPNMF with resting-state based brain parcellation (RS-parcellation)

The well above chance level (≈.30) adjusted rand index between the RS-parcellation and the OPNMF indicates that the spatial representations derived form OPNMF based on structural covariance converge well with the spatial representation derived from resting-state functional signal in health adults. Of note, the used RS-parcellations have both been extensively evaluated in their respective studies, namely with regards to stability and convergence with histological mapping and alternative parcellations (Bellec et al. 2010; Schaefer et al. 2017). Thus, the similarity between the brain partitions derived from OPNMF and the "optimized" RS-parcellation that we found in the current study allows us to assume that our OPNMF brain partitions have some biological validity. Similar observations have been reported recently by Sotiras et al. 2017 who showed that at low granularity (< 60), the components derived by OPNMF resembles previously evidenced functional brain networks. Together, these findings thus suggest that OPNMF of VBM data to some extend captures meaningful patterns of brain functional organization, both at the network and areal level.

While OPNMF-based factorization and RS-parcellation showing good convergence, they did not show a perfect agreement. This is in line with the few multi-modal mapping studies showing that brain maps from different features (such as structure and function) converge towards similar brain partition schemes, but also suggesting that each feature targeting a specific aspect of the brain tissue, each feature can capture an unique aspect of brain organization (Kelly et al. 2012; Genon et al. 2016; Glasser et al. 2016; Genon et al. 2017). In other words, different features (i.e. modalities) are to some extend sensitive to different aspects of brain organization (for a more detailed discussion see Eickhoff et al. 2017). From the perspective of data compression, the most efficient partitions should thus come from the same modality. And indeed, RS-parcellations provides more homogeneous parcels when assessing resting-state images than histologically defined brain regions (Craddock et al., 2012). This leaves the question, whether the amount of transferable information is still

sufficient for a useful representation. Our results also provided evidence that this is the case by showing that a more accurate age prediction model is built from VBM data when this data is compressed directly as compared to representing it based on a functional parcellation of the brain (Fig 5) even if the latter yields very good accuracies. Overall, in the context of multivariate pattern analysis, we suggest that brain parcellation derived from one modality is transferable to another modality for data reduction even though it does not reach within-modality performance.

4.4 Influence of different datasets on the OPNMF

Despite the brain topographical pattern of the OPNMF components derived from the two different datasets show similar convergence with the independent RS-parcellation and general good agreement between them, they are not perfectly similar (Fig. 1B & 1C; Fig S5), However, it has to be noted that the similarity between the factorizations derived from the two datasets has been measured at a level of granularity that does not favor reproducibility (even between datasets which are age, gender and site matched) according to previous work (Sotiras et al. 2015; Sotiras et al. 2017; Fig S5). As the granularity increases, the resolution of structural covariance increases resulting in a finer representation of covariance patterns, but that are, in turn, more influenced by covariance trends specific to the dataset used. Thus, we assumed that the difference in sample characteristics between the two cohorts could explain the slight differences in the brain topographical pattern of the factorizations.

4.5 Impact of granularity on age prediction

Importantly, our study showed that despite the fact that reproducibility may decrease at high level of granularity (Sotiras et al. 2017), prediction performance did increase as granularity increases (as previously suggested by Sotiras et al., 2015). Our results demonstrated that when the number of components reaches approximately 300-400, prediction accuracy remains largely stable, particularly when the factorization is derived from the same dataset. However, when the factorization is derived from an independent dataset, a somewhat higher granularity (i.e., a few more than 400 components) might be required to reach stability. At a level of granularity around 300-400 components, OPNMF seemed to factorize the entire voxel-wise data into efficient subdivisions, which allowed the LASSO regression model to capture only the relevant features (i.e., ~116 features when predicting MIXED sample and ~52 features when predicting 1000BRAINS sample) and ignore the unnecessary/noisy features relatively better than at coarser granularity. Our finding converges with the study of Franke et al. (2010), in which the data compression was performed using PCA. These authors found that the lowest mean absolute error of the prediction analysis was reached at around 350 components. Of note, this level of granularity (or factorization) seems also convergent with the range of subdivisions of the brain that emerged as stable in functional MRI data, which lies between 200 and 500 parcels (Tucholka et al. 2008; Thirion et al. 2014; Gordon et al. 2016, Schaefer et al., in press). We could assume that a lower level of

subdivision (i.e., $< \approx 200$ components) provides less homogeneous regions (i.e., regions mixing different functional and structural properties, cf. Eickhoff et al. 2017), while a higher level of subdivisions ($> \approx 500$ components) might spatially narrow the components but without importantly improving the homogeneity within regions. Thus, the current study suggests that a factorization of VBM data into 300 to 500 components optimally organizes voxel-wise structural data into homogeneous brain regions for age prediction.

4.6 Model validation within the same dataset using a cross-validation approach

In the model validation within the same dataset, our study showed that the performance of the brain age prediction using the framework of non-negative sparse coding (i.e., non-negative matrix factorization with LASSO regression model) was similar to the prediction accuracy found in previous studies (Franke et al. 2010; Liem et al. 2017). It is important to note that predicting the age of subjects compressed using the components derived from the same dataset violates the test set independence. Even though the subjects in the test dataset were separated from the training dataset at the prediction level, the used factorization reflects the best factorization of the entire dataset, including the test dataset (Yuan et al. 2007; Liu et al. 2010). In other words, the test dataset cannot be considered as strictly unseen because the test data has been "optimally spatially organized" with its own factorization scheme. Hence, performing the brain age estimations on the dataset compressed using the same dataset's factorization facilitates optimistic predictions. Therefore, in this study we compared the performance of the proposed prediction framework with this later over-optimal approach.

Our results demonstrated that the LASSO regression, when applied on the dataset compressed with components derived from an independent dataset, estimated the brain age with a precision comparable to that achieved when compressing the dataset with its own factorization. This finding confirms the previous literature arguing that the 'projectivity' of OPNMF allows the efficient transferability of the factorization onto a new unseen dataset (Yuan et al. 2007). Our results showed that the differences between the factorizations derived from the two datasets (cf. section 4.3) did not influence the prediction of brain age when transferring the components onto the other dataset. Furthermore, our supplementary results (Fig. S3) illustrated that the pattern of regions selected by the prediction approach remained similar when the factorization was derived from another dataset. That is, the prediction model recollected the same anatomical regions regardless of which factorization scheme was applied. Again, this pattern of findings converges with what has been previously observed in data reduction of fMRI data for subsequent functional connectivity analyses. Those parcellation studies have observed that at an optimal resolution, parcellation from one dataset can provide a relevant spatial representation of the functional signal in other datasets, despite the topographical pattern of the parcellation between the datasets being different (Bellec et al. 2010; Finn et al. 2015; Gordon et al. 2016). Similarly, OPNMF factorization based on a different dataset did not prevent an optimal compression of the data for age prediction or the

selection of the relevant (anatomical) features. Thus, overall, our results demonstrated that despite the fact that factorization results from different datasets may comprise slightly different spatial components, any of the stable factorizations offers an efficient data compression for prediction analyses.

4.7 Model validation on an independent dataset
In the context of dataset (in-) homogeneity from different protocols, we observed that the prediction model extrapolated quite well to an independent dataset (Fig 6B & Fig 7B: Top left & bottom left plots). Firstly, model trained on a highly heterogeneous dataset, better predicted the subjects' age in an independent dataset. Thus, our study supported Liem et al. 2017's recent suggestions that merging datasets from multiple protocols could avoid fitting the model to the characteristics of a particular scanner protocol. In other words, heterogeneous datasets allow the model to encounter a wider range of variations, helping it to disentangle non-relevant inter-individual variations from relevant variations for prediction. Surprisingly, model trained on single-site study also performed efficiently, when predicting the age of highly heterogeneous dataset (OldMIXED). To note, the single-site study consists of 693 subjects between 55 to 75 years (Fig 1A). Therefore, the model trained on this dataset encountered a wide range of variation at each age point. We suppose that this exposure to wide range of variation might have allowed overcoming the scanner effects with a robust regression model. Thus, we would recommend to train a given prediction model on a heterogeneous dataset (either with multi-sited examples or with multiple examples, ideally both) to ensure that true relevant variations are learned, which in turn may support good prediction performance. Importantly, the two cohorts also differed in their age distribution. Therefore, in addition to the generalizability over different protocols, these datasets also allowed us to evaluate the generalization of the prediction model over different age distributions (Fig 6B & Fig 7B: Top row plots). Not unexpectedly, models trained on restricted age range of the heterogeneous dataset (OldMIXED) provided better age prediction for test sample coming from an independent dataset within the age range of the training sample (1000BRAINS) when compared to the model trained on broader age range (MIXED). Again to be expected, models trained on narrow age range single-site study (1000BRAINS) failed to predict the age of subjects (MIXED) that were out of the training sample's age range. Together, these observations further confirmed the general recommendation for the prediction model to be trained on data comprising variations due to distinct parameters (such as the acquisition protocol and demographic characteristic). Despite the fact that this recommendation might sound trivial, it actually complements previous recommendations emphasizing the importance of sample size for good prediction performance (Varoquaux et al. 2012; Varoquaux et al. 2017), but further points out that, not the size per se matters, but the range of variations that are covered.

When applied in a clinical context, i.e. when evaluated on the ADNI dataset, the proposed

framework not only showed good age-prediction for the healthy subjects but in particular also captured premature aging in the context of MCI and dementia as indicated by positive BrainAGE scores (Fig 8). More specifically, the dramatic atrophy of AD patients was reflected by a mean BrainAGE score of almost 10 years, which is comparable to the findings of a previous study conducted by Franke et al. (2010). The sensitivity of our framework to brain structural changes in clinical populations was underscored by the likewise elevated BrainAGE for MCI patients, which was lower than for those with AD but still on average in the rage of 5-6 years, i.e., above the MAE in the population-based samples (Davatzikos et al. 2008; Franke and Gaser 2012). In other words, our framework accurately ranked the HC, MCI and AD groups with regards to their pathological and clinical progression from healthy to demented (considering MCI as a transitional stage between normal aging and dementia; Petersen, 2010). However, statistically discriminating those individuals among MCI patients who will progress towards Alzheimer's disease is a challenging issue (Davatzikos et al. 2009; Petersen 2010; Gaser et al. 2013; Moradi et al. 2015). While a classification approach could be more powerful for such purpose than age prediction (Franke and Gaser 2014; Wang et al. 2016; Beheshti et al. 2017), the latter could be combined with the former to quantify deviations from normal aging trajectories across clinical stages.

4.8 Brain Age estimation using our framework

Overall, our results demonstrate that models trained on highly heterogeneous life span sample (MIXED) can predict the age of any unseen subject with a precision of 6 years (irrespective of approach i.e., either on a cross-validation on MIXED dataset or on an independent dataset). Given the broad age range of the training sample (18 to 81 years), a precision of 6 years can be considered as a good performance from the technical side. Importantly, all previous brain age prediction studies likewise reported a precision of approximately 5 – 6 years in the context of life-span samples. The relationship between GMV and the chronological age is modulated by many factors (both environmental and genetic factors (Burgmans et al. 2009; Giedd et al. 2010; Harada et al. 2013; Luders et al. 2016; Cole et al. 2017)). When aiming to identify the relationship between brain structural pattern and age, those factors may introduce noise obscuring the systematic effects of age on brain structure. In addition to these factors, inclusion of participant with certain characteristics (such as, participants in younger age with unidentified subclinical brain alterations, or older adults representing above-normal (i.e. "super healthy") aged participants) might as well deviate the prediction model to capture the systematic effect of age on brain structure (Burgmans et al. 2009). Accordingly, these factors and the noise they introduce could account for the precision gap of 5 - 6 years in brain age prediction studies. That is, the limited precision of life-span age prediction may less relate to technical limitations but rather indicate that structural changes over a period of around 5 years are smaller than variations related to confounding factors that would represent "non-relevant" noise to the model. However, this hypothesis might not hold true for all life periods. For instance, one can

observe dramatic age-related structural changes in childhood (cf. Erus et al. 2015) the late life periods (cf. the higher precision of ~4 year MAE for 1000BRAINS or MIXED_55-90), while age-related grey matter changes could be relatively minor during early and middle adulthood (Schippling et al. 2017). Further examinations of these issues in future studies could provide better understanding of neurobiological aging. Nevertheless, in the scope of the present study, these confounding factors do not prevent our framework (combining OPNMF with sparse regression model) to accurately capture normal variations related to age and deviance from normal variations in clinical populations.

4.9 Identification of the regions influencing the prediction

As previously highlighted, our prediction model promotes interpretability of the prediction by inducing sparsity on the OPNMF representation. The OPNMF based low-rank approximations are more localized than the PCA components. Accordingly, using PCA-based compression of VBM data, Franke et al. 2012 observed that the features that contributed to the estimation of brain age constituted a pattern spread across the whole brain. In contrast, the framework examined in our study induces sparsity on the localized components, highlighting regions that are relevant for the prediction model. Our results revealed that the regions involved in the age prediction model in 1000BRAINS were sparser than those underlying the prediction of age in MIXED. In contrast, most of the brain regions (representing 73% of the total grey matter volume) seemed to underpin the prediction when the model was trained in the MIXED dataset (which covers the adult lifespan with subjects between 18 to 81 years old). Put simply, the model cannot be consistently restricted to a few regions for inferring subjects' age when the cohort covers the adult lifespan. Such a pattern could argue for a more complex pattern of grey matter variations across the whole adult life span than in the later life periods. Previous studies have demonstrated that many different patterns of changes occur across the adult life span in grey matter volume with notably some regions showing monotonic decrease of GM and other showing a clear inverted U-shape grey matter volume (GMV)-age relationships or a "delayed decline" (Ziegler et al. 2012; Douaud et al. 2014; Ziegler et al. 2014). Furthermore, as aforementioned, several factors may induce brain structural variations in the young and middle-aged adult brain, such as life style and environmental factors (Miller et al. 2016)) complicating the relationship between age and grey matter. In our study, in addition to the regions highlighted for age prediction in older sample, some regions, such as the amygdala, and the superior parietal lobule further contributed to age prediction when the model was trained on the young and middle age adults (MIXED) dataset. Interestingly, the amygdala is one region where GMV has been found to increase with age in relatively younger samples (8 to 30 years old; (Ostby et al. 2009)) and some authors have noted no age-related GMV changes in the amygdala in older samples (Good et al. 2001). Furthermore, structural covariance of the amygdala (with other brain regions) is known to be modulated by several factors such as gender (Mechelli 2005). Thus, we could assume that, in a prediction model mixing genders, the amygdala could be

selected as an indicator modulating the pattern of relationship between other brain regions and age, despite this region per se does not show a strong, linear and universal GMV decrease with age. Accordingly, when examining the pattern of association between GMV and age, we observed a mild general linear decline of grey matter volume with age, but with a high variance across the subjects in the MIXED sample (see Fig. S4) suggesting that very different age-related grey matter change patterns might be observed in this brain region. Such a pattern allows us to assume that the GMV in the amygdala, taken as isolated information, cannot significantly contribute to the age prediction, particularly in the case of older participants. In other words, we assume that the grey matter changes in the amygdala are diverse and occur mainly in the young and middle age adult lifespan rendering this specific region informative for predicting age in the whole adult life span sample in combination with information from other regions. However, on its own, this region would not be particularly informative for age prediction in older populations.

The superior parietal cortex is another example of regions contributing to the prediction analysis when the training sample consisted of young adults in addition to the older adults, but not when the model was trained on older adults only. Terribilli et al. (2011) conducted a study mainly focusing on young and middle age adults (18 to 50 years old), in which GMV of the lateral parietal cortex (i.e., supramarginal, angular and superior parietal cortex) exhibited a nonlinear relationship with age. The non-linear trend reported by the authors could be explained by a quadratic fit, that is, GMV followed a linear decline until the end of the fourth decade and then showed a mild increase. When examining the relationship between GMV and age in the superior parietal region in our study, we observed a similar trend (see Fig. S4), in which the mean GMV of the superior parietal region showed a sharp decrease until 40 years of age, but less pronounced change with age in later life. Thus, despite the fact that prediction models in general, (specially LASSO regression models) are inherently linear, identification of GMV in the superior parietal cortex as relevant for age prediction converges with previous data demonstrating that structural changes in this region occur mainly in the first decades of adult life, but not in periods later in life. Thus, visually examining the pattern of associations between GMV and age in regions contributing to the prediction in MIXED suggest that some regions may be informative for their relatively systematic changes in the first period of adult life (such as the superior parietal cortex) while others regions could contribute by introducing complementary information (such as the amygdala) despite not exhibiting a clear linear relationship with age across the sample.

The pattern of regions consistently contributing to the prediction in the older sample appeared more spatially specific. Many of the regions highlighted by these analyses such as the hippocampus, the temporo-occipital region and the medial superior frontal gyrus have been shown to be strongly affected by aging in the older life period and more specifically, to follow a strong linear decrease in this life period (after 40-50 years old; Raz et al. 2010a; Douaud et al. 2014). However, some other regions, such as the regions around the central

sulcus are not known to show systematic change with age in later life period. Thus, the pattern of regions contributing to the predictions in 1000BRAINS suggests that when the training sample is restricted to older populations, the model can be restricted to a few regions, whose grey matter volumes is systematically affected by the aging process in the later life period, as well as other regions that might not appear particularly informative form a neurobiological point of view but complement the information conveyed by the former regions.

Interpreting the multivariate brain pattern weighting in the prediction is usually not recommended (e.g. (Haufe et al. 2014) since the prediction is underlined by the combination of several element/feature (i.e. voxels in a voxel-wise representation of the data and components in the present study) and that the individual elements on themselves, taken in isolation, may not convey any neurobiological relevant information. However, we would argue that the relationship between the brain and the predicted variable should not be kept as a conceptually locked black box, that is, the multivariate aspect of the prediction does not imply that we should not at least try to understand why the given pattern is relevant for the model. As a metaphor, if a model uses the variable "number of children" and "country" for predicting the age of a person, obviously the variable "country" on its own is not informative for predicting the age of a person, in contrast, the number of children is partly informative. Hence, examining the combination of those two variables for predicting the age of a person can provide us more insight by suggesting that the relationship between age and the number of children is modulated by cultural factors. Similarly, the pattern of relationship between grey matter volume and age is assumed to be modulated by several factors, but whose influence remained relatively poorly understood. However, the current framework promoting spatially localized component as relevant features could help to explore this issue in future studies (such as how the complex pattern of structural variations in the amygdala influenced by gender can contribute to age prediction in healthy adults).

5 Conclusions and practical considerations for future studies

In conclusion, our study, which evaluated OPNMF-based compression of VBM data for age prediction in two different healthy adult cohorts, opens several new perspectives. First, we demonstrated that OPNMF compression allows age prediction with a precision that is well comparable to that achieved following PCA compression but yields substantially more interpretable results. It also outperformed an atlas-based approach based on resting-state whole-brain parcellation, even though the precision obtained by cross-model atlas based data compression is in itself remarkable. Considering the declining return of investment when going to higher granularities, we would thus suggest that OPNMF at a granularity of 300 and 500 components may provide the optimal data compression for age prediction.

While the exact OPNMF solution obviously depends on the examined sample, we here showed that prediction accuracies are basically uncompromised when employing a factorization derived from an independent dataset. That is, a factorization derived from one dataset can be used to efficiently compress VBM data of a second, independent dataset in a prediction framework. To note, the MIXED dataset used in the current study covers a wide range of variation over a broad age range while the (single-site) 1000BRAINS datasets can be assumed to capture structural covariance in older populations. Accordingly, the factorization derived from MIXED could be used for data compression in age prediction studies across adulthood whereas the factorization derived from 1000BRAINS may be particularly well suited for studying the aging brain. In addition to structural covariance-based factorization, our study offers robust prediction models trained on life span sample from heterogeneous sites (MIXED), an advantage on which future studies could capitalize to better understand the effects of different factors on the neurobiological aging.

Conflict of Interest

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

Ethical approval

The local ethics committees of the university hospital Aachen and Heinrich-Heine-University Düsseldorf have approved the original study protocol of the data used here. All the participants prior to the examination obtained informed consent. The current data were analyzed anonymously.

References

Ashburner J (2007) A fast diffeomorphic image registration algorithm. Neuroimage 38:95–113. doi: 10.1016/j.neuroimage.2007.07.007

Ashburner J, Yeung K, Ruzzo W, et al (1985) VBM Tutorial. J Classif 2:193–218. doi: 10.1007/BF01908075

Beheshti I, Demirel H, Matsuda H (2017) Classification of Alzheimer's disease and prediction of mild cognitive impairment-to-Alzheimer's conversion from structural magnetic resource imaging using feature ranking and a genetic algorithm. Comput Biol Med 83:109–119. doi: 10.1016/j.compbiomed.2017.02.011

Bellec P, Rosa-Neto P, Lyttelton OC, et al (2010) Multi-level bootstrap analysis of stable clusters in resting-state fMRI. Neuroimage 51:1126–1139. doi: 10.1016/j.neuroimage.2010.02.082

Boutsidis C, Gallopoulos E (2008) SVD based initialization: A head start for nonnegative matrix factorization. Pattern Recognit 41:1350–1362. doi: 10.1016/j.patcog.2007.09.010

Bunea F, She Y, Ombao H, et al (2011) Penalized least squares regression methods and applications to neuroimaging. Neuroimage 55:1519–1527. doi: 10.1016/j.neuroimage.2010.12.028

Burgmans S, van Boxtel MPJ, Vuurman EFPM, et al (2009) The prevalence of cortical gray matter atrophy may be overestimated in the healthy aging brain. Neuropsychology 23:541–550. doi: 10.1037/a0016161

Caspers S, Moebus S, Lux S, et al (2014) Studying variability in human brain aging in a population-based German cohort-rationale and design of 1000BRAINS. Front Aging Neurosci 6:1–14. doi: 10.3389/fnagi.2014.00149

Cole JH, Poudel RPK, Tsagkrasoulis D, et al (2017) Predicting brain age with deep learning from raw imaging data results in a reliable and heritable biomarker. Neuroimage 163:115–124. doi: 10.1016/j.neuroimage.2017.07.059

Craddock RC, James GA, Holtzheimer PE, et al (2012) A whole brain fMRI atlas generated via spatially constrained spectral clustering. Hum Brain Mapp 33:1914–1928. doi: 10.1002/hbm.21333

Davatzikos C (2016) Computational neuroanatomy using brain deformations: From brain parcellation to multivariate pattern analysis and machine learning. Med. Image Anal. 33:149–154.

Davatzikos C (2004) Why voxel-based morphometric analysis should be used with great caution when characterizing group differences. Neuroimage 23:17–20. doi: 10.1016/j.neuroimage.2004.05.010

Davatzikos C, Fan Y, Wu X, et al (2008) Detection of prodromal Alzheimer's disease via pattern classification of magnetic resonance imaging. Neurobiol Aging 29:514–523. doi: 10.1016/j.neurobiolaging.2006.11.010

Davatzikos C, Xu F, An Y, et al (2009) Longitudinal progression of Alzheimers-like patterns of atrophy in normal older adults: The SPARE-AD index. Brain 132:2026–2035. doi: 10.1093/brain/awp091

Douaud G, Groves AR, Tamnes CK, et al (2014) A common brain network links development, aging, and vulnerability to disease. Proc Natl Acad Sci U S A 111:17648–17653. doi: 10.1073/pnas.1410378111

Eickhoff SB, Constable RT, Yeo BTT (2017) Topographic organization of the cerebral

cortex and brain cartography. Neuroimage

Erus G, Battapady H, Satterthwaite TD, et al (2015) Imaging patterns of brain development and their relationship to cognition. Cereb Cortex 25:1676–1684. doi: 10.1093/cercor/bht425

Finn ES, Shen X, Scheinost D, et al (2015) Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity. Nat Neurosci 18:1–11. doi: 10.1038/nn.4135

Franke K, Gaser C (2012) Longitudinal changes in individual BrainAGE in healthy aging, mild cognitive impairment, and Alzheimer's disease. GeroPsych J Gerontopsychology Geriatr Psychiatry 25:235–245. doi: 10.1024/1662-9647/a000074

Franke K, Gaser C (2014) Dementia classification based on brain age estimation. Proc MICCAI Work Chall Comput Diagnosis Dement Based Struct MRI Data 48–54.

Franke K, Gaser C, Manor B, Novak V (2013) Advanced BrainAGE in older adults with type 2 diabetes mellitus. Front Aging Neurosci 5:1–9. doi: 10.3389/fnagi.2013.00090

Franke K, Luders E, May A, et al (2012) Brain maturation: Predicting individual BrainAGE in children and adolescents using structural MRI. Neuroimage 63:1305–1312. doi: 10.1016/j.neuroimage.2012.08.001

Franke K, Ristow M, Gaser C (2014) Gender-specific impact of personal health parameters on individual brain aging in cognitively unimpaired elderly subjects. Front Aging Neurosci 6:1–14. doi: 10.3389/fnagi.2014.00094

Franke K, Ziegler G, Klöppel S, Gaser C (2010) Estimating the age of healthy subjects from T1-weighted MRI scans using kernel methods: Exploring the influence of various parameters. Neuroimage 50:883–892. doi: 10.1016/j.neuroimage.2010.01.005

Gaser C, Franke K, Klöppel S, et al (2013) BrainAGE in Mild Cognitive Impaired Patients: Predicting the Conversion to Alzheimer's Disease. PLoS One. doi: 10.1371/journal.pone.0067346

Genon S, Li H, Fan L, et al (2017) The Right Dorsal Premotor Mosaic: Organization, Functions, and Connectivity. Cereb Cortex 27:2095–2110. doi: 10.1093/cercor/bhw065

Genon S, Reid A, Li H, et al (2016) The heterogeneity of the left dorsal premotor cortex evidenced by multimodal connectivity-based parcellation and functional characterization. Neuroimage

Giedd JN, Stockman M, Weddle C, et al (2010) Anatomic magnetic resonance imaging of the developing child and adolescent brain and effects of genetic variation. Neuropsychol. Rev. 20:349–361.

Glasser MF, Coalson TS, Robinson EC, et al (2016) A multi-modal parcellation of human cerebral cortex. Nature 536:171–178. doi: 10.1038/nature18933

Good CD, Johnsrude IS, Ashburner J, et al (2001) A voxel-based morphometric study of ageing in 465 normal adult human brains. Neuroimage 14:21–36. doi: 10.1006/nimg.2001.0786

Gordon EM, Laumann TO, Adeyemo B, et al (2016) Generation and Evaluation of a Cortical Area Parcellation from Resting-State Correlations. Cereb Cortex 26:288–303. doi:

10.1093/cercor/bhu239

Guyon I, Elisseeff A (2003) An Introduction to Variable and Feature Selection. J Mach Learn Res 3:1157–1182. doi: 10.1016/j.aca.2011.07.027

Harada CN, Natelson Love MC, Triebel KL (2013) Normal cognitive aging. Clin. Geriatr. Med. 29:737–752.

Haufe S, Meinecke F, G??rgen K, et al (2014) On the interpretation of weight vectors of linear models in multivariate neuroimaging. Neuroimage 87:96–110. doi: 10.1016/j.neuroimage.2013.10.067

Hua J, Tembe WD, Dougherty ER (2009) Performance of feature-selection methods in the classification of high-dimension data. Pattern Recognit 42:409–424. doi: 10.1016/j.patcog.2008.08.001

Hubert L, Arabie P (1985) Comparing partitions. J Classif 2:193–218. doi: 10.1007/BF01908075

Jolliffe IT (2002) Principal Component Analysis, Second Edition.

Kampa K, Mehta S, Chou CA, et al (2014) Sparse optimization in feature selection: Application in neuroimaging. J Glob Optim 59:439–457. doi: 10.1007/s10898-013-0134-2

Kelly C, Toro R, Di Martino A, et al (2012) A convergent functional architecture of the insula emerges across imaging modalities. Neuroimage 61:1129–1142. doi: 10.1016/j.neuroimage.2012.03.021

Koutsouleris N, Davatzikos C, Borgwardt S, et al (2014) Accelerated brain aging in schizophrenia and beyond: A neuroanatomical marker of psychiatric disorders. Schizophr Bull 40:1140–1153. doi: 10.1093/schbul/sbt142

Lakkaraju H, Bach SH, Jure L (2016) Interpretable Decision Sets: A Joint Framework for Description and Prediction. Kdd 2016:1675–1684. doi: 10.1145/2939672.2939874

Lee DD, Seung HS (1999) Learning the parts of objects by non-negative matrix factorization. Nature 401:788–91. doi: 10.1038/44565

Liem F, Varoquaux G, Kynast J, et al (2017) Predicting brain-age from multimodal imaging data captures cognitive impairment. Neuroimage 148:179–188. doi: 10.1016/j.neuroimage.2016.11.005

Liu X, Yan S, Jin H (2010) Projective nonnegative graph embedding. IEEE Trans Image Process 19:1126–1137. doi: 10.1109/TIP.2009.2039050

Luders E, Cherbuin N, Gaser C (2016) Estimating brain age using high-resolution pattern recognition: Younger brains in long-term meditation practitioners. Neuroimage 134:508–513. doi: 10.1016/j.neuroimage.2016.04.007

May A (2011) Experience-dependent structural plasticity in the adult human brain. Trends Cogn. Sci. 15:475–482.

Mechelli A (2005) Structural Covariance in the Human Cortex. J Neurosci 25:8303–8310. doi: 10.1523/JNEUROSCI.0357-05.2005

Miller KL, Alfaro-Almagro F, Bangerter NK, et al (2016) Multimodal population brain imaging in the UK Biobank prospective epidemiological study. Nat Neurosci 19:1523–1536.

doi: 10.1038/nn.4393

Moradi E, Pepe A, Gaser C, et al (2015) Machine learning framework for early MRI-based Alzheimer's conversion prediction in MCI subjects. Neuroimage 104:398–412. doi: 10.1016/j.neuroimage.2014.10.002

Mwangi B, Hasan KM, Soares JC (2013) Prediction of individual subject's age across the human lifespan using diffusion tensor imaging: A machine learning approach. Neuroimage 75:58–67. doi: 10.1016/j.neuroimage.2013.02.055

Mwangi B, Tian TS, Soares JC (2014) A review of feature reduction techniques in Neuroimaging. Neuroinformatics 12:229–244.

Ostby Y, Tamnes CK, Fjell AM, et al (2009) Heterogeneity in subcortical brain development: A structural magnetic resonance imaging study of brain maturation from 8 to 30 years. J Neurosci 29:11772–11782. doi: 10.1523/JNEUROSCI.1242-09.2009

Petersen RC (2010) Alzheimer's disease: progress in prediction. Lancet Neurol 9:4–5. doi: 10.1016/S1474-4422(09)70330-8

Raz N, Ghisletta P, Rodrigue KM, et al (2010) Trajectories of brain aging in middle-aged and older adults: Regional and individual differences. Neuroimage 51:501–511. doi: 10.1016/j.neuroimage.2010.03.020

Santos JM, Embrechts M (2009) On the use of the adjusted rand index as a metric for evaluating supervised classification. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). pp 175–184

Schaefer A, Kong R, Gordon EM, et al (2017) Local-Global Parcellation of the Human Cerebral Cortex from Intrinsic Functional Connectivity MRI. Cereb Cortex 1–20. doi: 10.1093/cercor/bhx179

Schippling S, Ostwaldt AC, Suppa P, et al (2017) Global and regional annual brain volume loss rates in physiological aging. J Neurol 264:520–528. doi: 10.1007/s00415-016-8374-y

Sotiras A, Resnick SM, Davatzikos C (2015) Finding imaging patterns of structural covariance via Non-Negative Matrix Factorization. Neuroimage 108:1–16. doi: 10.1016/j.neuroimage.2014.11.045

Sotiras A, Toledo JB, Gur RE, et al (2017) Patterns of coordinated cortical remodeling during adolescence and their associations with functional specialization and evolutionary expansion. Proc Natl Acad Sci 114:3527–3532. doi: 10.1073/pnas.1620928114

Terribilli D, Schaufelberger MS, Duran FLS, et al (2011) Age-related gray matter volume changes in the brain during non-elderly adulthood. Neurobiol Aging 32:354–368. doi: 10.1016/j.neurobiolaging.2009.02.008

Thirion B, Varoquaux G, Dohmatob E, Poline JB (2014) Which fMRI clustering gives good brain parcellations? Front Neurosci 8:1–13. doi: 10.3389/fnins.2014.00167

Tibshirani R (1996) Regression Selection and Shrinkage via the Lasso. J. R. Stat. Soc. B 58:267–288.

Tipping M (2001) Sparse Bayesian Learning and the Relevance Vector Mach. J Mach Learn

Res 1:211–244. doi: 10.1162/15324430152748236

Tipping ME, Faul AC (2003) Fast Marginal Likelihood Maximisation for Sparse Bayesian Models. Ninth Int Work Aritficial Intell Stat 1–13. doi: 10.1.1.165.4281

Tisserand DJ, Pruessner JC, Sanz Arigita EJ, et al (2002) Regional frontal cortical volumes decrease differentially in aging: An MRI study to compare volumetric approaches and voxel-based morphometry. Neuroimage 17:657–669. doi: 10.1016/S1053-8119(02)91173-0

Tucholka A, Thirion B, Perrot M, et al (2008) Probabilistic anatomo-functional parcellation of the cortex: How many regions? In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). pp 399–406

Varoquaux G, Gramfort A, Thirion B (2012) Small-sample brain mapping: sparse recovery on spatially correlated designs with randomization and clustering. Proc 29th Int Conf Mach Learn 1375–1382.

Varoquaux G, Raamana PR, Engemann DA, et al (2017) Assessing and tuning brain decoders: Cross-validation, caveats, and guidelines. Neuroimage 145:166–179. doi: 10.1016/j.neuroimage.2016.10.038

Wang P, Chen K, Yao L, et al (2016) Multimodal classification of mild cognitive impairment based on partial least squares. J Alzheimer's Dis 54:359–371. doi: 10.3233/JAD-160102

Yang Z, Oja E (2010) Linear and Nonlinear Projective Nonnegative Matrix Factorization. Trans Neur Netw 21:734–749. doi: 10.1109/TNN.2010.2041361

Yang Z, Yuan Z, Laaksonen J (2007) Projective Non-negative matrix factorization with application to facial image processing. Pattern Recognit 21:1353–1362.

Yuan Z, Yang Z, Oja E (2007) Projective Nonnegative Matrix Factorization : Sparseness , Orthogonality , and Clustering. Technology 1–14.

Zhang CH, Huang J (2008) The sparsity and bias of the lasso selection in high-dimensional linear regression. Ann Stat 36:1567–1594. doi: 10.1214/07-AOS520

Zhao P, Yu B (2006) On Model Selection Consistency of Lasso. J Mach Learn Res 7:2541–2563. doi: 10.1109/TIT.2006.883611

Ziegler G, Dahnke R, Jäncke L, et al (2012) Brain structural trajectories over the adult lifespan. Hum Brain Mapp 33:2377–2389. doi: 10.1002/hbm.21374

Ziegler G, Ridgway GR, Dahnke R, Gaser C (2014) Individualized Gaussian process-based prediction and detection of local and global gray matter abnormalities in elderly subjects. Neuroimage 97:333–348. doi: 10.1016/j.neuroimage.2014.04.018

Zou H, Hastie T (2005) Regularization and variable selection via the elastic-net. J R Stat Soc 67:301–320. doi: 10.1111/j.1467-9868.2005.00503.x

I. Supplementary methods

I.1. Scanning parameters:

The dataset termed "1000BRAINS" was obtained from a unique data collection [Caspers et al., 2014] of 693 healthy older adults (age: 55-75 years) that were scanned on a 3T scanner. Scanning parameters for the 1000BRAINS dataset were as given: Repetition time = 2.25 s, Echo time = 3.03 ms, Inversion time= 900ms, Field of view = 256 x 256 mm^2, flip angle = 9^0, voxel resolution = 1 x 1 x 1 mm^3.

Whereas, the dataset termed "Mixed" comprised 1084 healthy adults (age: 18-81 years) that were collected by merging samples of healthy adults from multiple sites. Scanning parameters for the individual samples that were merged to form the Mixed dataset were as following.

Dataset-1: No of subjects = 91, Repetition time = 2.25 s, Echo time = 3.03 ms, flip angle = 9^0, voxel resolution = 1 x 1 x 1 mm^3, 3T TimTrio Siemens scanner.

Dataset-2: No of subjects = 83, Repetition time = 2.4 s, Echo time = 2.14 ms, flip angle = 8^0, voxel resolution = 0.7 x 0.7 x 0.7 mm^3, 3T Skyra Siemens scanner.

Dataset-3: No of subjects = 306, Repetition time = 1.9 s, Echo time = 2.52 ms, flip angle = 9^0, voxel resolution = 1 x 1 x 1 mm^3, 3T TimTrio Siemens scanner.

Dataset-4: No of subjects = 127, Repetition time = 2.5 s, Echo time = 3.5 ms, flip angle = 8^0, voxel resolution = 1 x 1 x 1 mm^3, 3T TimTrio Siemens scanner.

Dataset-5: No of subjects = 50, Repetition time = 2.25 s, Echo time = 3.26 ms, flip angle = 9^0, voxel resolution = 1 x 1 x 1 mm^3, 3T TimTrio Siemens scanner.

Dataset-6: No of subjects = 26, Repetition time = 3 s, Echo time = 4 ms, flip angle = 8^0, voxel resolution = 1 x 1 x 1 mm^3, Philips Achieva 3T scanner.

Dataset-7: No of subjects = 42, Repetition time = 2.25 s, Echo time = 3.82 ms, flip angle = 9^0, voxel resolution = 1.04 x 1.04 x 1 mm^3, 3T Siemens scanner.

Dataset-8: No of subjects = 78, Repetition time = 2.3 s, Echo time = 2.92 ms, flip angle = 9^0, voxel resolution = 1 x 1 x 1.1 mm^3, 3T TimTrio Siemens scanner.

Dataset-9: No of subjects = 13, Repetition time = 2.3 s, Echo time = 3.03 ms, flip angle = 9^0, voxel resolution = 1 x 1 x 1 mm^3, 3T TimTrio Siemens scanner.

Dataset-10: No of subjects = 17, Repetition time = 1.9 s, Echo time = 2.52 ms, flip angle = 9^0, voxel resolution = 1 x 1 x 1.1 mm^3, 3T Magnetom Prismatit Siemens scanner.

Dataset-11: No of subjects = 72, Repetition time = 2.53 s, Echo time = 1.64 ms, flip angle = 7^0, voxel resolution = 1 x 1 x 1 mm^3, 3T TimTrio Siemens scanner.

Dataset-12: No of subjects = 35, Repetition time = 2.25 s, Echo time = 3.26 ms, flip angle = 9^0, voxel resolution = 1 x 1 x 1 mm^3, 3T TimTrio Siemens scanner.

Dataset-13: No of subjects = 31, Repetition time = 2.4 s, Echo time = 4.6 ms, flip angle = 30^0, voxel resolution = 1 x 1 x 1 mm^3, PHILIPS Intera Achieva 3T scanner.

Dataset-14: No of subjects = 16, Repetition time = 10 ms, Echo time = 4.6 ms, flip angle = 8^0, voxel resolution = 1 x 1 x 1 mm^3, 3T IRM Philips scanner.

Dataset-15: No of subjects = 77, Repetition time = 7.2 s, Echo time = 3.3 ms, flip angle = 8^0, voxel resolution = 0.89 x 0.89 x 0.9 mm^3, PHILIPS Achieva 3T scanner.

Dataset-16: No of subjects = 20, Repetition time = 9.86 s, Echo time = 4.6 ms, flip angle = 8^0, voxel resolution = 0.875 x 0.875 x 1 mm^3, PHILIPS Achieva 3T scanner.

I.2. ADNI dataset

For the clinical validation of our framework, we used data from the "ADNI" (Alzheimer's Disease Neuroimaging Initiative) database (www.loni.ucla.edu/ADNI) including the subjects that was scanned on a 3T scanner. The entire ADNI sample collection consists of three different phases i.e., ADNI-1, ADNI-GO and ADNI-2, with various modifications in the study design across phases. Hence, 3-0T MRI image acquisition scheme of ADNI-2 was modified according to the upgraded systems over the period of time. The subjects included in the current study were obtained from both ADNI-1 and ADNI-2 datasets. Scanning parameters for the ADNI-1 dataset were as follow: Repetition time = 0.65 s, Echo time = min full, Field of view = 256 x 256 mm^2, flip angle = 8^0, slice thickness = 1.2mm and scanning parameters for the ADNI-2 dataset were as follow: Repetition time = 0.4 s, Echo time = min full, Field of view = 256 x 256 mm^2, flip angle = 11^0, slice thickness = 1.2mm. General inclusion criteria and classification of the participants are described here: https://adni.loni.usc.edu/wp-content/uploads/2008/07/adni2-procedures-manual.pdf. Briefly, cognitively healthy participants do not report any specific memory complaints, neither show objective memory dysfunction. MCI, have both a subjective and an objective memory concern measured by education adjusted scores on Wechsler Memory Scale Logical Memory II, but an absence of significant levels of impairment in other cognitive domains, essentially preserved activities of daily living, and an absence of dementia. Finally, AD subjects who meet NINCDS/ADRDA criteria for probable AD were selected.

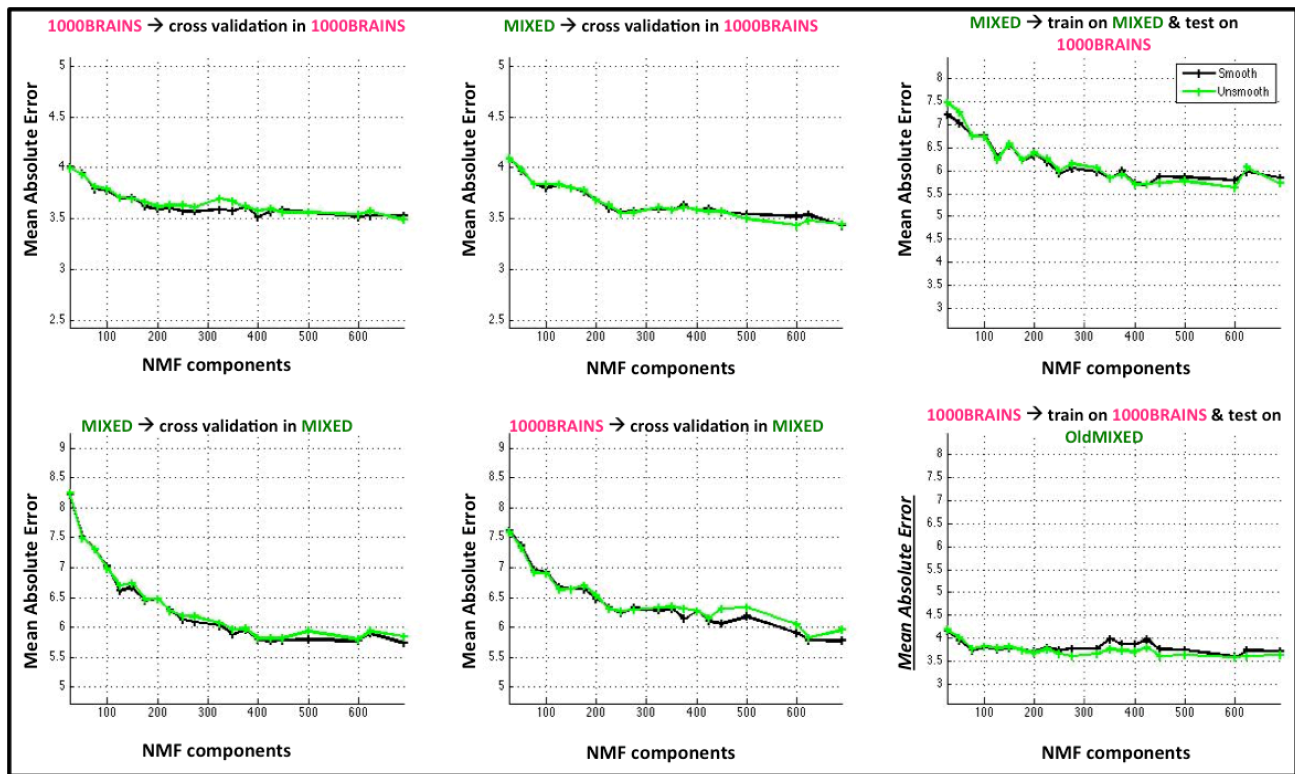## II. Supplementary results

Fig S1. Illustration of the mean absolute error at different levels of granularity using smoothed and unsmooth VBM data
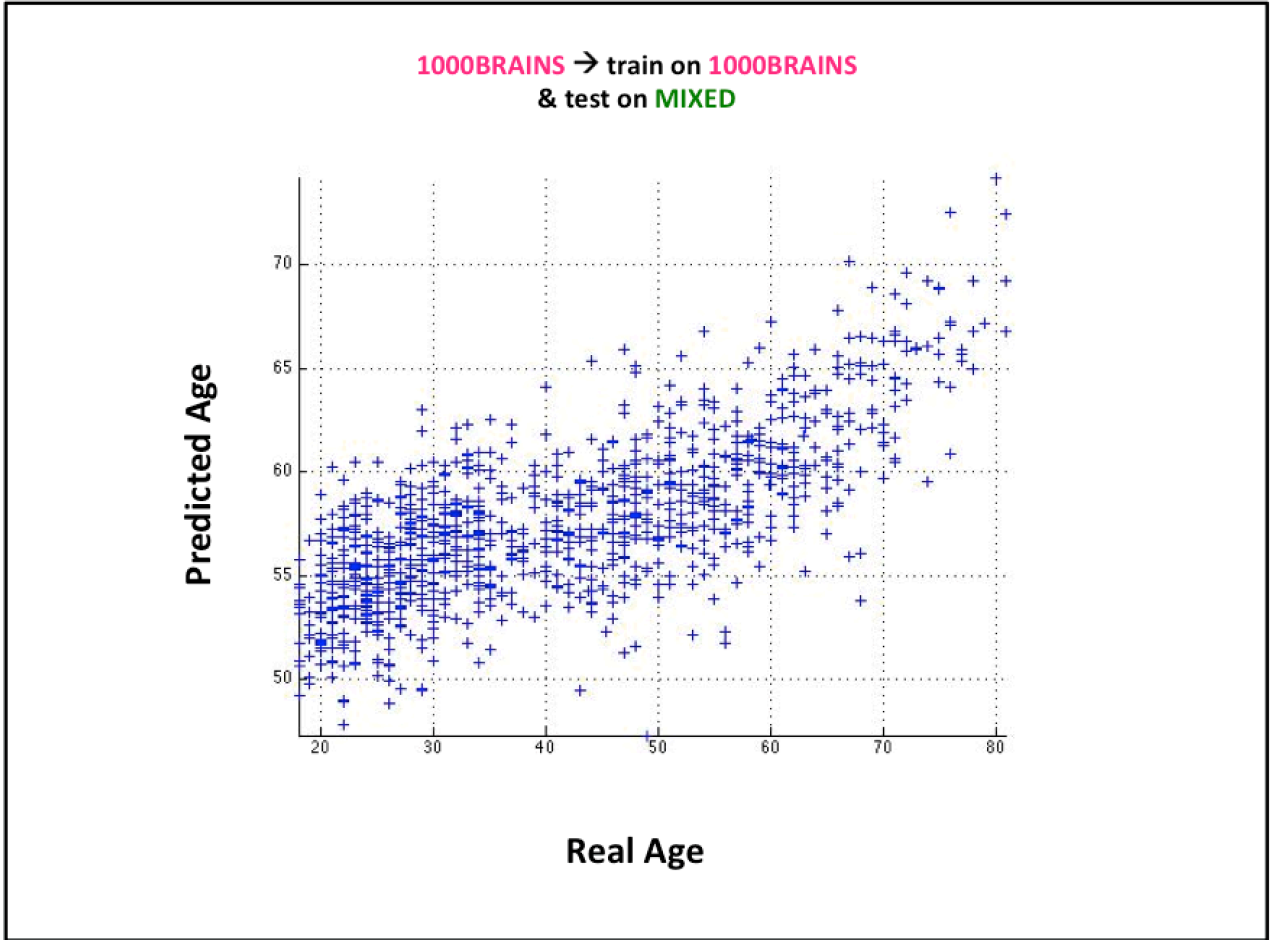
Fig S2. Illustration of the mean absolute error of the prediction model approach: 1000BRAINS → train on 1000BRAINS & test on MIXED: Training the prediction model on 1000BRAINS dataset projected on to its own factorization and later tested on MIXED dataset
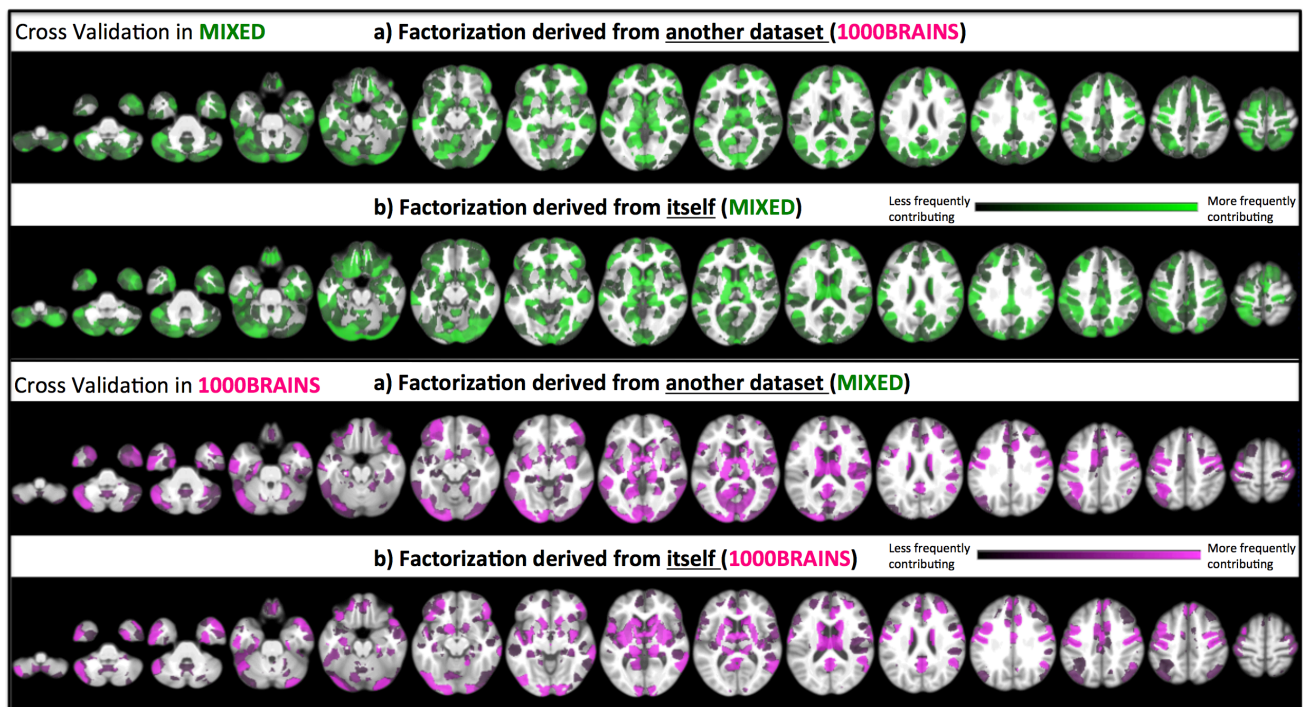
Fig S3. Summary map of the regions that contributed in the prediction analysis when performing 10-fold cross-validation and compressing the dataset using the components derived from its own dataset or their respective other group's dataset.

A) MIXED dataset compressed, a) with its own factorization, b) projected onto the factorization derived from 1000BRAINS.

B) 1000BRAINS dataset compressed, a) with its own factorization, b) projected onto factorization derived from MIXED.
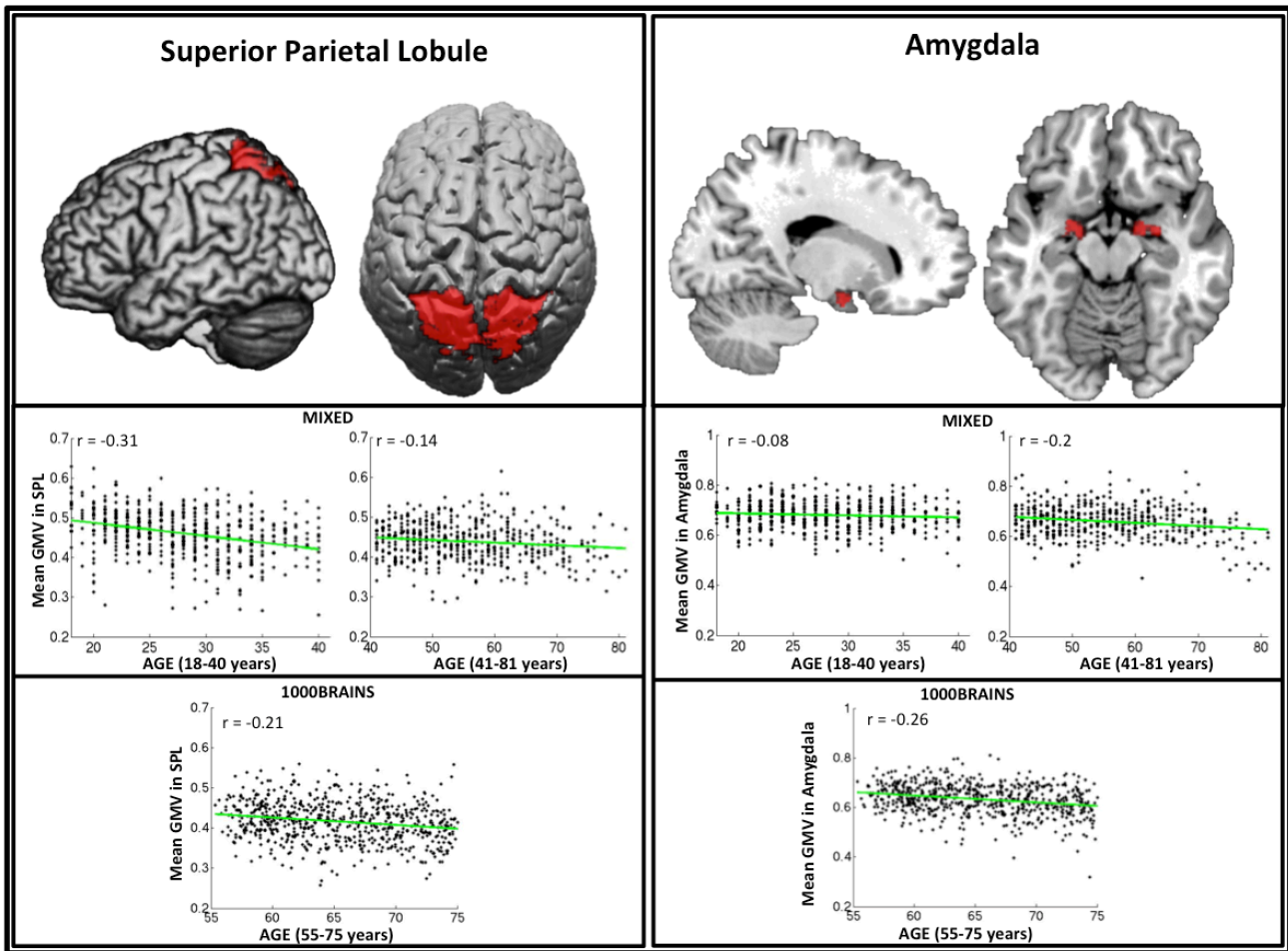
Fig S4. Scatter plots of the mean grey matter volume (GMV) of superior parietal lobule (SPL; top left) and amygdala (top right) across distinct age groups in the MIXED dataset (middle panel) and the full age range of the 1000BRAINS dataset (bottom panel). These regions were highlighted in the age prediction when the model was trained on the MIXED dataset, but were not selected when the model was trained on the 1000BRAINS dataset. The age range (y axis) on the scatter plots was displayed separately for the early adulthood (18-40 years) and the second life period (41-81) in the MIXED dataset illustrating different trends across the two life periods. The reported r values suggest a moderate negative correlation between superior parietal GMV and age in early adulthood and a weaker correlation in later life periods. In contrast, amygdala GMV show no GMV decrease with age in the early adulthood.
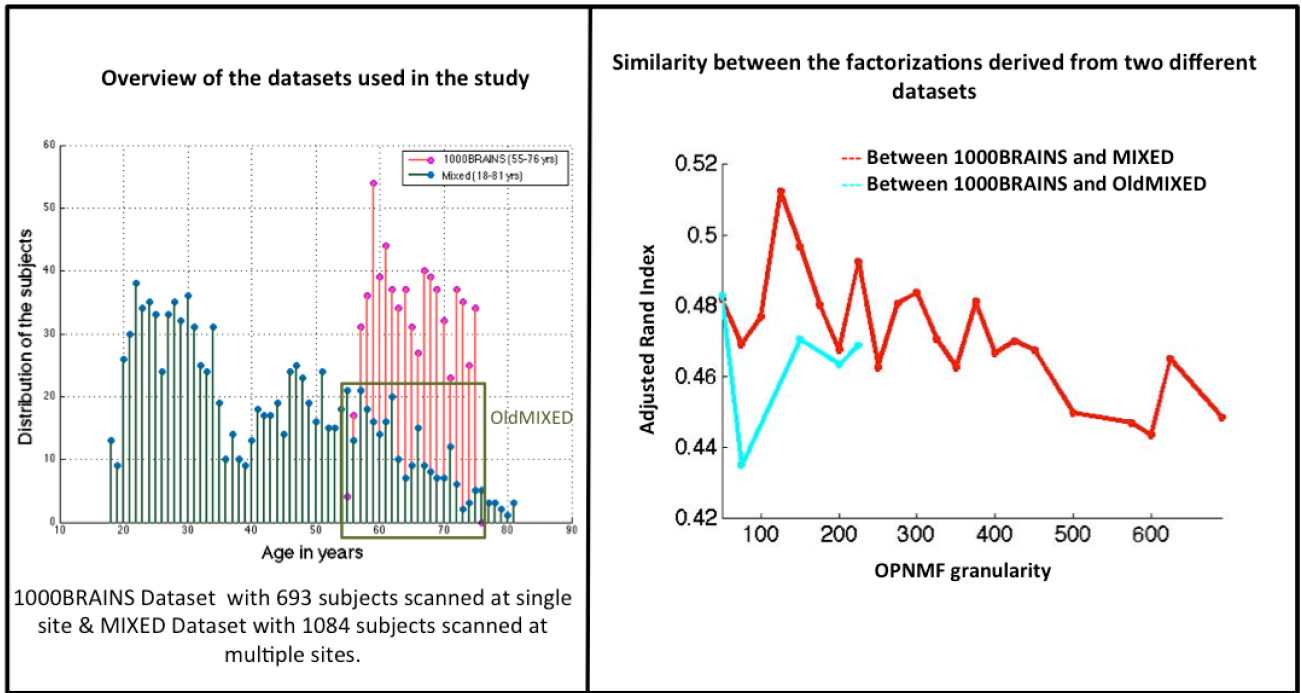
Fig S5. Similarity between factorizations from two datasets measured with adjusted Rand Index. Red: 1000BRAINS and MIXED; Cyan: 1000BRAINS and OldMIXED. Note: due to the small size of the OldMIXED (n = 239), granularity in this sample has been limited to 225 components.
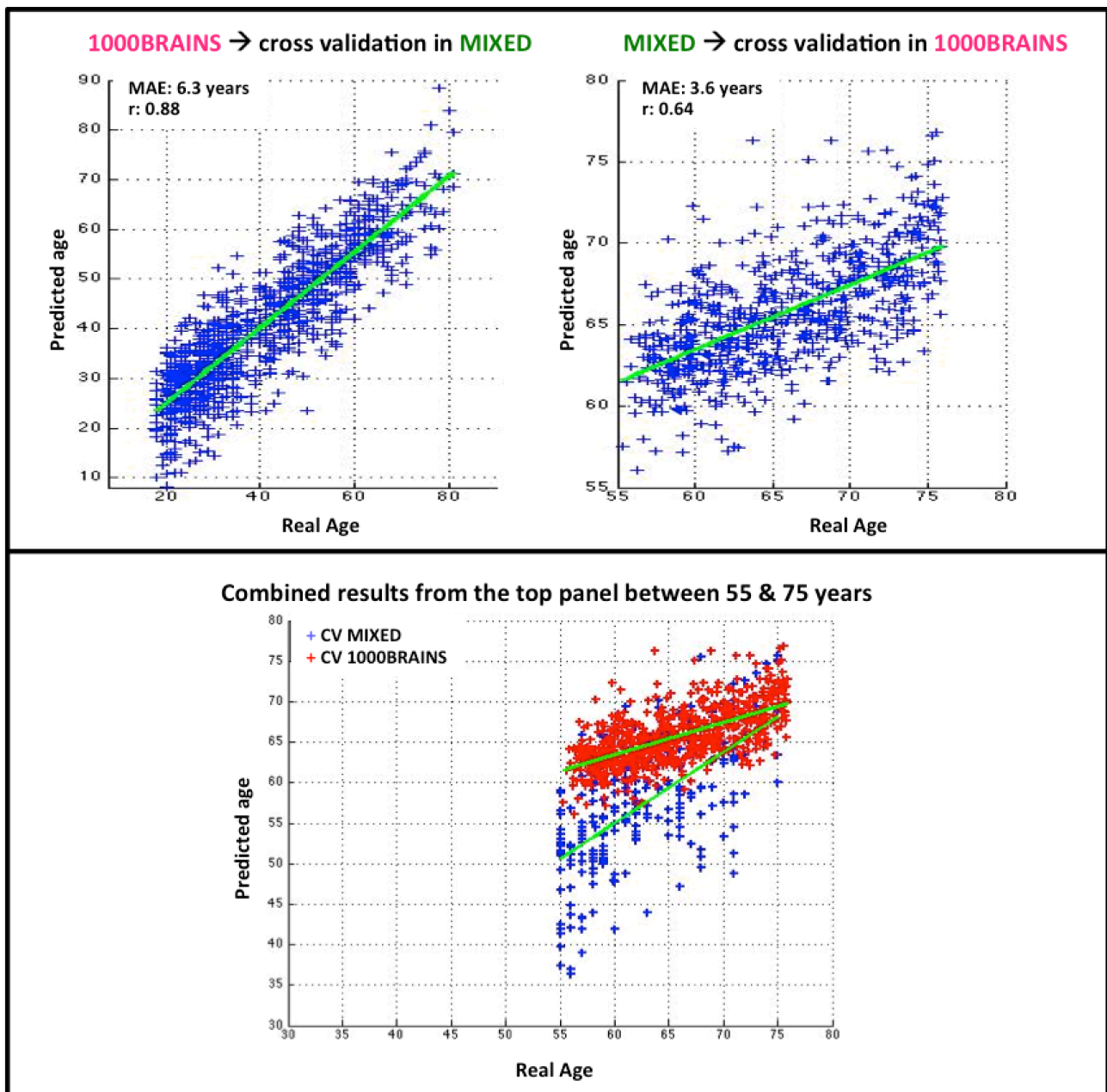
Fig S6. Predicted age plotted against chronological age, with fitted regression lines. In the bottom panel, both individual top plots have been combined together for subjects aged between 55 and 75.
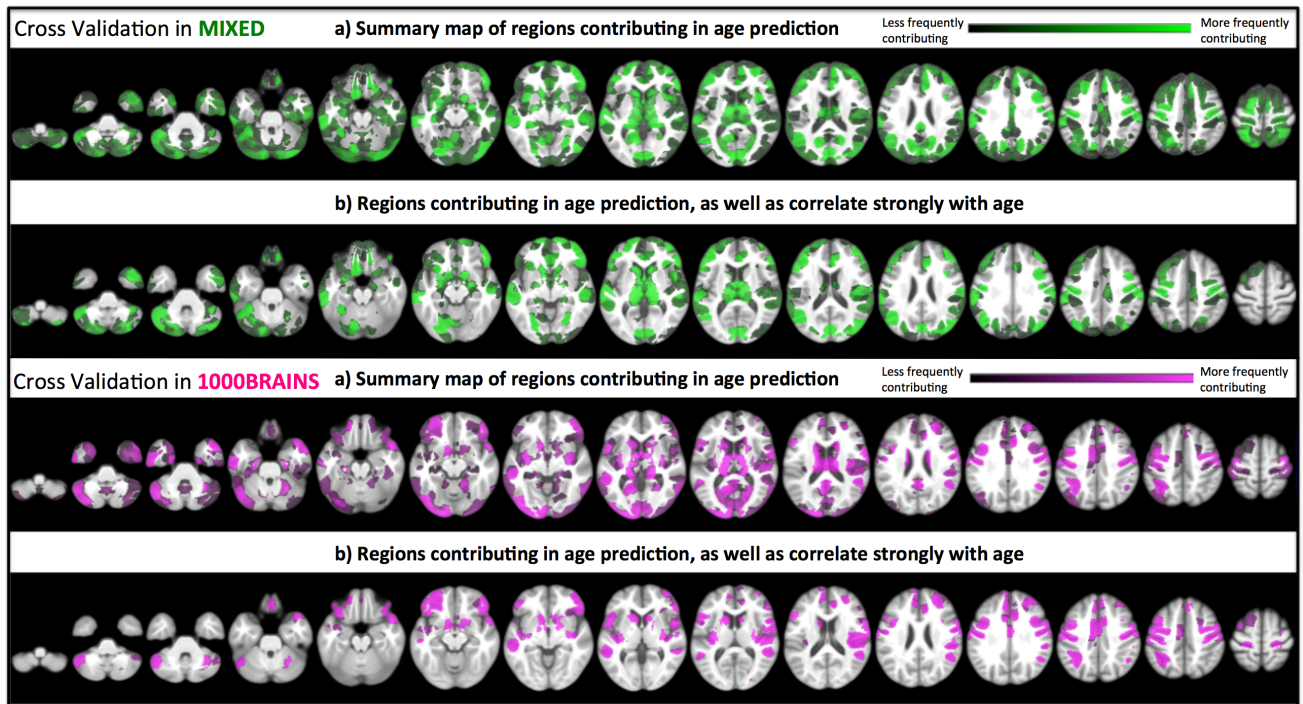
Fig S7. Summary map of the regions that contributed in the prediction analysis, when performing the cross-validation approach in either of the datasets (i.e., MIXED and 1000BRAINS). a) Summary map of the regions that contributed in the prediction analysis, b) Regions whose loading coefficients showed high correlations with age (among the regions contributing in age prediction)

Table S1: Mean absolute error averaged across different levels of granularity

Components derived	Training	Test	Whole sample	Males	Females
1000BRAINS	1000BRAINS	1000BRAINS	3.6	3.7	3.5
	MIXED	MIXED	6.1	6	6.3
	1000BRAINS	OldMIXED	3.7	3.7	3.7
MIXED	1000BRAINS	1000BRAINS	3.6	3.7	3.6
	MIXED	MIXED	6.3	6	6.5
	MIXED	1000BRAINS	6.1	6	6.2
	OldMIXED	1000BRAINS	3.8	3.8	3.8
Raw VBM	1000BRAINS	1000BRAINS	3.4	3.4	3.3
	MIXED	MIXED	5	4.9	5.1

Table S2: Correlations (<u>predicted age with real age</u>) averaged across different levels of granularity

Components derived	Training	Test	Whole sample	Males	Females
1000BRAINS	1000BRAINS	1000BRAINS	0.63	0.63	0.62
	MIXED	MIXED	0.87	0.88	0.86
	1000BRAINS	OldMIXED	0.53	0.53	0.53
MIXED	1000BRAINS	1000BRAINS	0.62	0.62	0.61
	MIXED	MIXED	0.86	0.87	0.85
	MIXED	1000BRAINS	0.55	0.55	0.54
	OldMIXED	1000BRAINS	0.54	0.54	0.54
Raw VBM	1000BRAINS	1000BRAINS	0.69	0.68	0.69
	MIXED	MIXED	0.91	0.91	0.91

Table S3: Computational time for each assessed prediction approach with 100 repetitions (in hours)

OPNMF from	Cross-validation approaches				Out of Sample approach		
	1000BRAINS	MIXED	1000BRAINS	MIXED	MIXED	1000BRAINS	MIXED
Granularity	CV 1000BRAINS		CV MIXED		Train MIXED Test 1000BRAINS	Train 1000BRAINS Test OldMIXED	Train OldMIXED Test 1000BRAINS
25	0.25	0.22	0.21	0.21	0.02	0.03	0.01
50	0.39	0.47	0.44	0.38	0.03	0.03	0.03
75	0.57	0.78	0.57	0.53	0.04	0.04	0.04
100	1.09	1.31	1.59	1.16	0.09	0.99	0.10
125	1.88	1.92	2.29	1.88	0.15	0.12	0.23
150	2.21	2.22	2.73	2.02	0.13	0.16	0.34
175	4.43	3.07	3.99	3.30	0.22	0.30	0.92
200	5.39	3.94	4.94	4.31	0.28	0.34	0.67
225	8.79	5.03	6.46	6.25	0.41	0.55	0.29
250	8.46	6.44	8.07	6.94	0.41	0.51	0.25
275	10.91	7.83	10.41	8.96	0.53	0.76	0.27
300	20.36	8.21	13.07	11.55	0.57	1.44	0.24
325	30.79	12.29	18.08	22.84	0.79	2.10	0.26

350	31.09	17.99	19.50	31.16	1.24	2.06	0.26
375	58.08	21.64	29.39	40.62	1.59	3.99	0.29
400	60.23	27.17	30.70	54.32	1.96	4.58	0.30
425	86.04	36.04	39.61	90.69	2.16	5.38	0.31
450	111.19	51.88	48.85	118.91	3.35	7.49	0.31
500	236.95	115.08	139.27	275.45	7.83	15.58	0.33
525	242.68	110.21	136.26	288.07	6.98	13.27	0.31
550	320.67	133.47	167.08	252.93	6.12	15.62	0.31
575	249.60	155.10	169.46	240.85	9.82	14.50	0.35
600	272.75	139.47	179.44	217.41	8.98	11.82	0.30
625	26.48	155.89	210.86	21.04	9.55	1.50	0.33
650	25	159.82	190.28	21.15	10.55	1.38	0.29
690	27.8	215.78	220.66	22.45	9.50	1.52	0.32

Table S4: Mean absolute error averaged across different levels of granularity

Components derived	Train	Test	Raw predictions		Adjusted predictions	
			Unsmoothed data	Smoothed data	Unsmoothed data	Smoothed data
1000BRAINS	*1000BRAINS*	*1000BRAINS*	3.6	3.6	4.6	4.5
	MIXED	MIXED	6.3	6.2	6.9	6.7
	1000BRAINS	MIXED	18.8	18	12.7	12
Mixed	1000BRAINS	1000BRAINS	3.6	3.6	4.6	4.55
	MIXED	*MIXED*	6.4	6.3	7	6.9
	MIXED	1000BRAINS	6.1	6.1	8.1	8
Raw VBM	1000BRAINS	1000BRAINS	3.4	3.3	3.3	3.4
	MIXED	MIXED	4.92	5	4.86	5

General Discussion

1 Extraction of functional connectivity measure in a reliable fashion

The first study was dedicated to investigate the influence of various confound removal procedures and signal extraction approaches on the reliability of the functional connectivity scores in a priori defined canonical networks. The results related to the extraction of the signal demonstrated that implementation of the grey matter masking based on the group-averaged GM probabilities improved the reliability of the connectivity measures for larger, a priori defined clusters, while no grey matter masking seems favorable when using smaller, spherical ROIs. In details, when using larger, a priori defined clusters, restricting the extraction of the signal from the grey matter masking based on individual GM probabilities has shown to improve the reliability at the connection level (i.e., within-subject connectivity). In turn, group grey matter masking enhances the reliability of the connectivity measures at group level. This finding argues for best fit of spatial correspondence reflecting individual subject's anatomy to produce reliable measure within a single subject. When performing a multivariate approaches, connectivity measures characterizing dynamics of the spatial information on individual basis is essential to improve the performance of the analysis (Davatzikos 2016). Therefore, individual grey matter masking seems advisable for optimizing the performance of a machine-learning algorithm on the larger, a priori defined clusters.

Our results further demonstrated that the confound removal strategy which retains the most variance is the most reliable strategy. One assumption behind this outcome was the that highest retained variance presumably included structured noise, such as reliable nuisance signals induced by physiological processes hence contributing to the reliability. In line with this assumption, Birn et al. 2014 reported a reduction in the test-retest reliability particularly after elimination of artifacts induced by physiological processes (i.e., cardiac and respiratory processes) and suggested reducing the spatially structured fluctuations in order to improve the validity of the functional connectivity measures. Therefore, we infer that structured noise caused by the systematic and cyclic physiological processes might have contributed to the highest reliability for the connectivity measures with no confound removal. Nevertheless, influence of physiological noise correction on the reliability of functional connectivity measures was not explicitly investigated in our study 1, due to the lack of physiological recordings of parameters such as heartbeat and breathing. In turn, we investigated PCA de-noising which has been introduced by Behzadi et al. (2007) as an alternative de-noising strategies and was shown to effectively eliminate the influence of physiologically induced artifacts (Chai et al. 2012). Even though, our main findings indicated that PCA denoising reduced the reliability of the connectivity measures, implementation of PCA denoising led high reliability at the connection level (i.e., within subject connectivity; RoCO) for within

network connections (only). Importantly, Shirer et al. in a study investigating the test-retest reliability of the connectivity measures has reported an increase in the group discriminability (i.e., an increase in the machine learning classification accuracy between Alzheimer's patients and healthy controls) after regressing out the PCA components. Together, we could assume that PCA denoising has a potential to reduce the intra-subject variability, preferably by eliminating the structured noise. Nevertheless, it has to be noted that PCA denoising aggressively remove the variance, which could make the time series reach a flat line, hence, resulting in a higher reliability at the connection level reported in our study. Thus, excessive removal of variance using PCA denoising might not be advisable. In sum, removal of physiological noise allow us to improve the validity of the measure by diminishing the detrimental structured noise, however, PCA denoising might not be the suitable approach. Furthermore, we note that physiological recordings are rarely acquired in standard (clinical) resting-state acquisitions, and hence remain challenging to precisely model the physiological noise regressors. Recently, Salimi-Khorshidi et al. 2014 proposed an ICA based FIX denoising strategy, which performs an automatic identification of confounds addressing the structured noise and regress them out. However, FIX denoising approaches require effective individual segmentation from high-resolution T1 images, which were not available for the data used in study 1. Hence, unfortunately, our study 1 failed to explicitly investigate the impact of the FIX denoising strategy on the reliability of the connectivity measures. However, we assume this automated denoising approach could efficiently model the true noise components in the existing data, given the main strength of this automated denoising, to detect the noise components based on machine-learning approach using data-driven feature selection strategy. Therefore, along with the confound removal suggested in our study 1 (i.e., 24 motion regressors and mean signal from WM and CSF), we suggest to improve the validity by eliminating the accurately modeled physiologically induced nuisance regressors either by using pre-acquired physiological recordings or by an automated data driven approaches (such as ICA-based FIX denoising).

2 Sparsity induced feature reduction

In this section, we discuss the results from study 4 based on the anatomical data, focusing on the performance and interpretability of a sparse supervised regression model employed with three different form of the data representation, i.e., 1) high-dimensional Voxel Based Morphometric (VBM) data without any implementation of data reduction procedure (uncompressed data), 2) Non-negative matrix factorization (NMF) based VBM data reflecting a sparse decomposition of the data and 3) PCA based VBM data representing a non-sparse data reduction procedure (implemented by the previous studies).

An interesting outcome of study 4 is that prediction accuracy remains analogous employing the three forms of data representation (i.e., using the full (uncompressed) VBM data, sparse

NMF compressed data and non-sparse PCA decomposition from the previous studies). However, investigation of underlying pattern within the prediction revealed that isolated voxels scatter all over the brain were identified as relevant features by the sparse regression model on the voxel-wise (uncompressed) data rather than forming clusters in definite regions. In addition, as reported in the Table 1, on an average each regression model selected 236 and 508 voxels (i.e., features) for 1000BRAINS and MIXED, nevertheless, only 5 and 24 features (for 1000BRAINS and MIXED) among these selected features remain consistent over (only) 50 percent of the regression models. This observation clearly suggest that high dimensional voxel wise data due to the highly correlated features allowed several potential models accurately fitting on the same data. This problem reveals lack of stability in the prediction model due to the multicollinearity of the data (Zhao and Yu 2006), which implicitly questions the interpretability of the pattern associated with the features contributed in the prediction model. In addition to this issue, as suggested by Wang et al. 2010, the small sample size relative to the high dimensional voxel-wise data (commonly known as small-n-large-p issue) can challenge the computation of a regression model. Thus, implementation of dimensionality reduction procedure, prior to the sparse supervised machine-learning algorithm is recommended to avoid the above-mentioned issues (namely, multicollinearity and small-n-large-p issues) and enhance the accuracy and stability of the results. In support to the above discussion, our study 2 focusing on functional MRI data actually had reported an increase in the classification accuracy for a non-sparse supervised machine-learning algorithm using a meta-analytically derived networks (domain knowledge based data reduction method) compared to the whole brain functional connectivity networks. This result further supports the argument of implementing a feature reduction method before the supervised learning to reduce the detrimental impact of the redundant, as well as noisy features.

Although, efficient feature reduction or dimensionality reduction has been the biggest concern in the field (Hua et al. 2009; Chu et al. 2012), deeper understanding on the pattern associated with the regions contributed within the supervised learning remains an open question which deserve equal attention (Franke et al. 2012). Thus, our project intended to evaluate an alternative dimensionality reduction method (i.e., NMF) that converts the high dimensional voxel-wise data into biologically plausible spatial units of the brain. As stated above, accuracy of the brain age prediction using NMF (sparse) data compression favorably compares with previous reports employing PCA (non-sparse) data compression. Furthermore, as reported in the study 4, NMF attempted to capture the most influential spatial information among the entire high-dimensional representation (Kim and Tidor 2003; Sotiras et al. 2015). Therefore, NMF has a potential for being less affected by the measurement and processing artifacts. Most importantly, study 4 demonstrated that implementation of a sparse regression model on to the sparse decomposition method (in particular, NMF) allowed us to investigate the underlying pattern contributing to the

prediction analysis. As enlightened by Davatzikos 2016, and supported by Wang et al. 2010, categorizing spatially consistent brain voxels (i.e., voxels of similar signals) across individuals into each component allowed us to partly articulate which groups of components (or features) provide discriminative information about the variable of interest. Thus, deeper and better understanding of the underlying biological processes can be achieved by Generative-discriminative methods (i.e., combination of sparse unsupervised feature reduction with a sparse supervised learning (Davatzikos 2016)), particularly when generative method promotes biologically plausible factorization of brain data.

In addition, results from study 3 focusing on the functional MRI data implementing domain knowledge based data reduction procedure (i.e., a-priori defined meta analytical networks) allowed us to strengthen our assumptions. Here, implementation of meta-analytically derived networks on a sparse regression model offered a deeper interpretation of the neurobiological underpinnings of personality traits. Our a-priori defined meta-analytically derived networks have a great potential in improving the signal to noise ratio of the feature space, which is a representation of relatively small number of brain regions implicated in the functional processing of the variable of interest. Relatedly, Chu et al. 2012 has evaluated the influence of different sample sizes along with different feature selection methods on classification accuracy and demonstrated that feature selection procedure capitalizing on the prior knowledge (with a prerequisite of high reliability) resulted in an improvement in the performance of the regression model. Most importantly, even though, study 2 has employed non-sparse supervised regression model (meaning when there isn't any explicit feature selection method implemented), application of the prior knowledge based functionally relevant features (i.e., meta-analytically defined networks as a feature reduction step) has partly allowed us to interpret the underlying pattern across the neurodegenerative biological processes. Altogether, feature reduction with an inherent or indirectly induced sparsity is highly beneficial for understanding the pattern associated within the analysis.

Table 1: proportion of features contributing in the prediction analysis on high dimensional voxel wise

No of features (i.e., voxels)	Cross validation in 1000BRAINS		Cross validation in MIXED	
	Average No of features selected per model (among 75 models)	Features contributed consistently in 50 % of the models	Average No of features selected per model (among 75 models)	Features contributed consistently in 50 % of the models
344383	236	5	508	24T

3 To what extent do we need to reduce the features?

Furthermore, study 4 demonstrated that prediction accuracy improves with an increase in the level of granularity. To note, sparse and non-sparse unsupervised models encounter a downside of the procedure, which is user-specified definition of the level of granularity as a prerequisite (Mwangi et al., 2015). In this context, study 4 suggested an optimal choice of granularity at 300 to 400 components using NMF (sparse representation), particularly for the VBM data in a prediction's purpose. Consistent with our finding, Franke et al. 2010 reported a comparable level of factorization (i.e., 350) using PCA (non-sparse representation). Importantly, the level of granularity reflecting a stable subdivision of the brain in functional MRI data also converged with our findings (i.e., ranging between 200 and 500 parcels (Tucholka et al. 2008; Thirion et al. 2014; Gordon et al. 2016)). Thus, granularities between 300 and 400 components might optimally compresses the voxel-wise structural MRI data into homogeneous spatial components. Importantly, this level of compression could also favorably subdivide the brain into stable functionally homogenous when considering fMRI data (Schaefer et al., 2017). Of note, sparse descriptive (supervised) models applied on the low rank approximations has a potential to enforce an automated selection of as many features as required by the regression model to provide relevant discriminative information about the variable of interest (Tipping and Faul 2003; Zhao and Yu 2006). Nevertheless, study 4 demonstrated that accuracy of the descriptive models improves with an increase in the level of granularity for the generative models. In addition, Table 2 depicts that the relevant number of features consistently contributing among 95 % of the prediction models

reaches a plateau after a certain level of granularity. These observations strengthen our assumption that determining a precise granularity for the generative reduction step is a fundamental requirement to achieve better precision of the descriptive models. Therefore, even though the descriptive models can perform an automated feature selection, an optimal a-priori definition of granularity for the feature reduction is imperative.

Table 2: proportion of features contributed in the prediction analysis on NMF based VBM data

Total No of features (Granularity)	Comp FZJ Cross Mixed		Comp Mixed Cross FZJ	
	Average No of features selected per model	Features contributed consistently in 95 % of the models	Average No of features selected per model	Features contributed consistently in 95 % of the models
25	24	21	24	20
50	48	40	39	27
75	65	51	42	29
100	91	75	49	25
125	113	86	74	35
150	124	83	90	40
175	136	88	98	53
200	142	92	94	46
225	149	91	106	55
250	175	113	103	49
275	186	115	106	45
300	181	114	111	51
325	195	118	111	50
350	206	120	117	50
375	220	124	127	46
400	215	113	118	53
425	210	105	121	46
475	237	110	138	52
500	225	114	128	43
525	255	126	147	57
550	233	110	142	57

575	229	111	131	50
600	252	122	130	52
625	258	109	155	48
650	275	126	143	56
690	248	111	167	64

4 Generalizability

It has been emphasized in the previous section that outcomes of the generative models (i.e., models enforcing unsupervised learning of the data by inferring a hidden pattern of the data across the observations) can impact the accuracy of the descriptive models (i.e., models learning the data in a supervised fashion). However, generative or unsupervised models based solutions aim to discover groups that share similar information across the individuals (i.e., observations) without any pre-defined regulations, which might raise a question of generalizability. In support to this assumption, study 4 reported a slightly different NMF estimated from two different datasets with distinct parameters (such as the acquisition protocol and demographic characteristic). As discussed in the Study 4, these differences might have been driven from the inter-individual variations in the topographical patterns of the brain. However, the most important observation from the study 4 is that factorization derived from one dataset can be efficiently implemented to reduce the dimensions of an independent dataset in a (sparse supervised learning) prediction framework. This observation advocates for the generalizability of the unsupervised learning (particularly, NMF) for an unseen data, which is not only independent but also, differs with acquisition protocol and demographic characteristics. In contrast, Study 4 also demonstrated that supervised learning or descriptive models failed to generalize for the test sample with target variable out of the range of variations covered by the training samples. However, supervised prediction model trained on data comprising largely heterogeneous variations due to divergent and wide range of parameters (such as acquisition protocol and demographic characteristic), efficiently performed the predictions on an independent dataset. Therefore, our results emphasize the importance of big sample sizes with large and heterogeneous range of target variables for effectively improving the generalizability of the supervised learning or descriptive modeling.

5 Summary

In the current thesis, best combination of confound removal strategies and signal extraction approaches to compute reliable functional connectivity measures were examined along with the evaluation of various data reduction methods. Firstly, to summarize the outcomes of the test-retest reliability study, when implementing meta-analytic approach as a feature reduction method, extraction of the signal by employing the grey matter masking has shown to improve the reliability of the connectivity measures for large, a priori defined clusters, in turn there is no requirement of implementing a grey matter masking for smaller, spherical ROIs. Along with the 24 motion regressors and mean signal from WM and CSF, physiologically induced noise regression should also be given adequate consideration. Now coming to the summary of the data reduction strategies, prior to the supervised machine-learning application, reduction of high dimensional data into fewer dimensions seems promising. In addition, feature reduction procedure with an inherent or indirectly induced sparsity resulting biologically plausible reduced units is highly beneficial to interpret the underlying pattern of the biological processes. Definition of a precise optimal latent dimension of the entire voxel wise representation is crucially required to attain reliably meaningful results from generative-descriptive methods. Generalization of the unsupervised learning (particularly, NNMF) for an unseen data, which is not only independent but also, differs with acquisition protocol and demographic characteristics seems achievable. Finally, we proposed that not only the size of the observation (i.e., sample size), but also heterogeneity on the variations associated to the range of target variables is crucial for enhancing the generalizability of the machine learning approach.

6 Future work

While this thesis has demonstrated the potential of sparse feature reduction techniques to better understand the prediction analysis, many opportunities for extending the scope of this thesis remain open. The following ideas could be explored in future studies.

1.   Given the main objective of this dissertation was to investigate approaches that would promote biologically meaningful compression of the high-dimensional space in the context of multivariate analysis. A linear regression model was preferred for the sake of comparison and complementation of previous studies, as most of the previous studies have implemented linear regression models. Despite the potential benefits of the linear regression models, it is indeed interesting to implement a non-linear regression model, which might increase the accuracy of the multivariate approaches, an issue that should be further investigated in future work.

2.	Even though, application of prior knowledge based meta-analytically defined networks has allowed us to efficiently perform the prediction or classification analysis. It is crucial to perceive for the global effects of the multivariate approaches based on whole-brain FC patterns. Therefore, in future studies, it will be interesting to evaluate the performance of multivariate approaches on whole brain FC pattern compressed using Non-negative matrix factorization (yielding spatially localized sub units).

3.	One major aspect that could be investigated as an extension to this dissertation is to address the curse of dimensionality at multimodal level. The fundamental issue underlying this aspect is the data fusion, which requires compressed representation in identical forms for all the modalities. Thus, either standardizing the low-rank approximations at each modality or implementing the same data compression method on all the modalities could support to explore this issue in future studies.

References

Behzadi Y, Restom K, Liau J, Liu TT (2007) A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. Neuroimage 37:90–101. doi: 10.1016/j.neuroimage.2007.04.042

Birn RM, Cornejo MD, Molloy EK, et al (2014) The influence of physiological noise correction on test-retest reliability of resting-state functional connectivity. Brain Connect 4:511–22. doi: 10.1089/brain.2014.0284

Chai XJ, Castañán AN, Öngür D, Whitfield-Gabrieli S (2012) Anticorrelations in resting state networks without global signal regression. Neuroimage 59:1420–1428. doi: 10.1016/j.neuroimage.2011.08.048

Chu C, Hsu AL, Chou KH, et al (2012) Does feature selection improve classification accuracy? Impact of sample size and feature selection on classification using anatomical magnetic resonance images. Neuroimage 60:59–70. doi: 10.1016/j.neuroimage.2011.11.066

Davatzikos C (2016) Computational neuroanatomy using brain deformations: From brain parcellation to multivariate pattern analysis and machine learning. Med. Image Anal. 33:149–154.

Franke K, Luders E, May A, et al (2012) Brain maturation: Predicting individual BrainAGE in children and adolescents using structural MRI. Neuroimage 63:1305–1312. doi: 10.1016/j.neuroimage.2012.08.001

Franke K, Ziegler G, Klöppel S, Gaser C (2010) Estimating the age of healthy subjects from T1-weighted MRI scans using kernel methods: Exploring the influence of various parameters. Neuroimage 50:883–892. doi: 10.1016/j.neuroimage.2010.01.005

Gordon EM, Laumann TO, Adeyemo B, et al (2016) Generation and Evaluation of a Cortical Area Parcellation from Resting-State Correlations. Cereb Cortex 26:288–303. doi: 10.1093/cercor/bhu239

Hua J, Tembe WD, Dougherty ER (2009) Performance of feature-selection methods in the classification of high-dimension data. Pattern Recognit 42:409–424. doi: 10.1016/j.patcog.2008.08.001

Kim PM, Tidor B (2003) Subsystem Identification Through Dimensionality Reduction of Subsystem Identification Through Dimensionality Reduction of Large-Scale Gene Expression Data. Genome Res 13:1706–1718. doi: 10.1101/gr.903503

Salimi-Khorshidi G, Douaud G, Beckmann CF, et al (2014) Automatic denoising of functional MRI data: Combining independent component analysis and hierarchical fusion of classifiers. Neuroimage 90:449–468. doi: 10.1016/j.neuroimage.2013.11.046

Sotiras A, Resnick SM, Davatzikos C (2015) Finding imaging patterns of structural covariance via Non-Negative Matrix Factorization. Neuroimage 108:1–16. doi: 10.1016/j.neuroimage.2014.11.045

Thirion B, Varoquaux G, Dohmatob E, Poline JB (2014) Which fMRI clustering gives good brain parcellations? Front Neurosci 8:1–13. doi: 10.3389/fnins.2014.00167

Tipping ME, Faul AC (2003) Fast Marginal Likelihood Maximisation for Sparse Bayesian Models. Ninth Int Work Aritficial Intell Stat 1–13. doi: 10.1.1.165.4281

Tucholka A, Thirion B, Perrot M, et al (2008) Probabilistic anatomo-functional parcellation of the cortex: How many regions? In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). pp 399–406

Wang Y, Fan Y, Bhatt P, Davatzikos C (2010) High-dimensional pattern regression using machine learning: From medical images to continuous clinical variables. Neuroimage 50:1519–1535. doi: 10.1016/j.neuroimage.2009.12.092

Zhao P, Yu B (2006) On Model Selection Consistency of Lasso. J Mach Learn Res 7:2541–2563. doi: 10.1109/TIT.2006.883611

ACKNOWLEDGMENTS

The journey till now would not have been possible without the guidance and support of several individuals who contributed and extended their valuable assistance that I am extremely grateful for.

First and foremost, I am deeply indebted to Simon Eickhoff for providing me an opportunity to work in his group. Throughout the journey, he has been an incredible mentor, an untiring teacher, and a determined collaborator. He has been a very kindhearted boss and well supported me during the hard times of my personal life for which i remain grateful for the rest of my life. His creativity and dedication for science has always taught me so much about being an excellent researcher. This work would never be the same without his persistent support, advice, inspiration and patience.

I extend my gratitude to Holger Schwender for his willingness to supervise and assess my dissertation. I was especially glad to be able to rely on your expertise regarding the administrative and scientific requirements of dissertations at the Heinrich Heine University, as these can seriously impede for accomplishing a doctoral degree

I give my most sincere gratitude to Sarah Genon, for her amazing kindness, incredible patience, innumerable time and effort invested to help me in multiple ways. She constantly encouraged and motivated me and always respected my crazy idea, without her it would have been tremendously exhausting. She has been extremely patient in answering several questions and always providing the most accurate strategies for communicating my thoughts and ideas. Most importantly, she has been an adorable friend, sister, guide, what not, everything to me during these years.

I am also grateful to Felix Hoffstaedter, Andrew Reid, Veronika Meuller (Moni), Edna Cieslik, and Robert Langner for so many insightful discussions, guidance to handle the administrative and computational hurdles, and also for their great company and interaction. It has been a true privilege to work with them. I would also like to thank all my research collaborators, colleagues and co-authors of the four studies reported in this dissertation for their time and effort invested with me, and their very helpful suggestions and comments while working on my thesis.

To my awesome co-PhD students, you are my strength and weakness, Alessandra nostro, Rachel Pläschke, Julia Camilleri, I would never be able to reach the end without your company. I have learned a lot from each of you "My Girls". I will always cherish your caring gestures until the end of time. Also, Anna Plachti, Julia Amunts, and Laura muzzarelli, your company and friendship is deeply appreciated, even though the times we spent were rather

short, they made me leave a very special place for you in my life. Altogether, our lunches, coffee, walks around the lake, blackboard times, bus journeys and laughing out loud moments really make me emotionally low to say 'this phase' has came to an end: thank you all for all the valuable memories. I know there are many more, and I will personally thank you.

Most importantly, I am thankful to my family, especially my lovely parents and my amazing sister for their unconditional love, care, support and patience. To my sweethearts, Lileesha and Chaitra, for being incredibly patient and tolerant, for being the best companion one can imagine. To all my closest friends, including company from 'SB11', I am thankful for sharing a lot of beautiful moments, interesting conversations, and incredible support during all these tough years. Last but not least, to my wonderful husband, for his love, understanding, and patience during this educational adventure.