Energetic cost of functional connectivity and effect of region specific regression on functional connectivity: Implications for resting-state fMRI

Behavioral and Cognitive Neuroscience, C-track, Major Thesis

July, 2016

Author:

Reinder A.J.A. Vos de Wael
UNIVERSITY OF GRONINGEN

Supervised by:

D. S. Fahmeed Hyder
YALE UNIVERSITY

Garth J. Thompson
YALE UNIVERSITY

Remco J. Renken
UNIVERSITY OF GRONINGEN

CONTENTS

1 Foreword 2

2 Introduction 3

3 Effects of region specific regression on resting-state networks 5

3.1 Introduction 5

3.2 Methods 6

3.2.1 Data Acquisition 6

3.2.2 Data Analysis 6

3.2.3 Ethical statement 7

3.3 Results 7

3.4 Discussion 9

4 Metabolic state alters the cost of functional connectivity 11

4.1 Introduction 11

4.2 Methods 11

4.2.1 Data Collection 11

4.2.2 Preprocessing and network creation 12

4.2.3 Data Analysis 13

4.3 Results 14

4.4 Discussion 15

5 Conclusion 16

6 Acknowledgements 16

7 References 17

8 Figures 19

9 Tables 27

1 FOREWORD

This thesis presents the results of my postgraduate associateship at Yale University. The original intent of my project was to research the effects of global signal regression (GSR) on brain connectivity. However, enroute to that long-term goal, my research also yielded interesting results interim that are not directly relevant to GSR. As a consequence, I have two distinct projects that are not easily presented together. Hence I have split this thesis into two chapters. The first chapter deals with the effects of gray matter, white matter, and global signal regression on brain connectivity, the second chapter deals with the relationship between the energy demand of spontaneous fluctuation in blood oxygen level dependent (BOLD). Each chapter has its own methods, results, and discussion. Both these chapters deal with analysis methods of fMRI data (amplitude of low frequency BOLD signal fluctuations and signal regression). There is some overlap, but as the methods are quite different in each chapter separating the projects provides a clearer story.

2 INTRODUCTION

Functional magnetic resonance imaging (fMRI) measures differences in blood oxygenation, which is often used as an indicator for neural activity. However, if this assumption is to hold true, then some mechanistic link between blood oxygenation and electric activity must exist. Neurons require energy for non-signaling processes (e.g. cell repair) and signaling processes (e.g. neural firing). The majority of energy expenditure of neurons is used for signaling, and because glucose oxidation is the most efficient mode of yielding large numbers of ATP, as such neural signaling is closely linked with oxidative metabolism [1-4]. If the fMRI signal can be interpreted in terms of metabolism, then this would provide a mechanistic link between fMRI signal and neural activity.

Glucose consumption in any cell, including neurons, breaks glucose down to either lactate, or water and carbon dioxide. First, glucose is broken down to pyruvate in a process called glycolysis. Glycolysis produces two ATP. Pyruvate can be converted to lactate, or further broken down in the Krebs cycle producing 36 ATP in the process. The breakdown of pyruvate through the Krebs cycle requires $6O_2$ whereas the conversion to lactate requires no oxygen at all. Hence, if the cerebral metabolic rate of glucose (CMR_{glc}) and oxygen (CMR_{O2}) are known, then the oxygen to glucose index (OGI) can be calculated as follows [5]:

$$OGI = \frac{CMR_{O2}}{CMR_{glc}}.$$

The OGI represents the ratio of glycolysis versus oxidative phosphorylation; OGI = 6 means that all glucose is oxidized (36 ATP/glucose), and OGI = 0 means that no glucose is oxidized (2 ATP/glucose). In the human cerebrum, OGI is approximately 5.5 [6] meaning that approximately 92% of glucose is oxidized. As there is a direct link between oxygen consumption and metabolism, and since blood oxygenation affects the blood oxygen level dependent (BOLD) signal, CMR_{O2} could provide a link between metabolism and the BOLD signal. Indeed, CMR_{O2} together with cerebral blood flow (CBF) and cerebral blood volume (CBV) determine the BOLD signal, or conversely CBF, CBV, and BOLD can be

used to estimate CMR_{O2} [7]. Moreover, through CMR_{O2} there is a link between the BOLD signal and the CMR_{glc}.

A simple manner of creating different metabolic states in human subjects' brains is asking them to keep their eyes closed or open; whilst subjects' eyes are open the metabolic activity in their brain increases [8]. Furthermore, in areas with more metabolic activity, the amplitude of low frequency fluctuations (ALFF) of the BOLD signal is increased [9]. This implies that there is a fractal structure in the BOLD signal i.e. the amplitude of the signal increases with the baseline of the signal (Figure 1).

The first chapter of this thesis concerns the effect of choice of region specific regression in preprocessing on the resulting BOLD signal. Different forms of region specific regression will be described and applied to several open-access resting-state fMRI (R-fMRI) datasets. It will be shown that the choice of region specific regression matters little, especially when compared with the choice between any regression or no regression. Furthermore, it will be shown that a region specific regression may hide artefacts in the data and that post-regression correlation maps look similar across datasets. The second chapter of this thesis studies the relationship between CMR_{glc} and BOLD amplitude in two different metabolic states. It will be shown that a linear relationship between metabolic activity and BOLD amplitude exists across metabolic states, however the coupling between them changes depending on the metabolic state.

3 EFFECTS OF REGION SPECIFIC REGRESSION ON RESTING-STATE NETWORKS

3.1 Introduction

During pre-processing of resting state fMRI data, a region specific regression is often performed. In signal regression the signal from "nuisance regions" and/or the motion parameters are regressed from all voxels. This regression is done to remove fluctuations that are unlikely to originate from physiological processes related to neural activity [10]. Common regions of interest for regression are the whole brain (a.k.a. global signal), white matter, and cerebral spinal fluid.

Global signal regression (GSR), has faced some controversy over the past few years. In a GSR a subject's mean time-series is regressed from all voxels' time-series. The underlying assumption is that the global signal is not involved in specific regional correlations and it reveals the anti-correlated default-mode network (DMN) and task positive network (TPN) [10]. However, GSR introduces anti-correlations by mathematical necessity [11], it removes metabolic information [12] and the global signal is correlated with local field potentials [13]. On the other hand, GSR is effective at reducing motion artefacts [14]. Due to these issues, GSR is a controversial method in fMRI preprocessing.

An alternative to GSR is white matter signal regression (WMSR). Whilst many researchers apply WMSR regressions in addition to or instead of WMSR [10, 15-18], I am unaware of an existing study that validates WMSR's efficacy for analyzing r-fMRI data. However, WMSR and GSR both reduce the metabolic component of the BOLD signal [12]. Thus, whilst WMSR does not create anti-correlations by mathematical necessity, it may act similar to GSR.

In this chapter of the thesis the differences between different types of signal regression are analyzed in terms of their effect on functional connectivity and motion correction.

3.2 Methods

3.2.1 Data Acquisition

Datasets for this experiment were downloaded from the 1000 functional human connectome project [19] (see table 1 for details). These datasets were selected such that the there was a spread of scanning parameters (e.g. gender, field strength, geographical location). Only the first twenty subjects were used for the Beijing-Zang dataset, only the female subjects were used in the Saint Louis dataset, and only the first functional sessions were used from the New Haven B dataset. Subjects were excluded if any abnormalities were mentioned in the documentation (Ontario: 2 subjects missing anatomical scans; Munich: 1 subject had fewer time-points in the functional scan). Two subjects (Orangeburg, Saint Louis) were excluded because registration failed.

3.2.2 Data Analysis

Standard preprocessing was performed as follows, but the preprocessing was stopped after region specific regression. This was done to prevent interaction effects between regressions and further preprocessing steps.

All R-fMRI datasets except for Ontario were slice-timing corrected, and all datasets were motion corrected. Ontario was not slice-timing corrected as it used spiral acquisition. Individual subjects' data were segmented to produce masks of gray matter, white matter, cerebrospinal fluid, and the whole-brain. R-fMRI data were registered to each subject's gray matter mask. The data in anatomical space was registered to the MNI template (2mm, isotropic) and blurred with a Gaussian filter (full width half maximum 8mm, size 6mm). To test the effects of signal regression, the following regressions were performed separately: global signal regression, gray matter signal regression, eroded white matter signal regression, and no regression. Global and eroded white matter signal regression were chosen as they are commonly used in the literature [10, 15-18], gray matter signal regression was chosen because in other R-fMRI data (data not shown) we found the gray matter signal to be nearly identical to the global signal. Regression signals were calculated by averaging the time-series within a whole

brain mask, a gray matter mask, or an eroded white matter mask. The white matter erosion removed all voxels with at least one side facing a non-white matter voxel thereby reducing the number of partial volume voxels. All regressions were performed with a least-squares linear regression.

Blurred (not regressed) R-fMRI data in MNI space of all subjects in all datasets and the Group ICA of fMRI toolbox (GIFT, mialab.mrn.org/software/gift/) [20] were used to generate 20 networks. Two of the ICA networks, the DMN and TPN, were identified by visual inspection.

All significance testing presented here was corrected for multiple comparison with the Benjamini-Hochburg procedure for controlling the false discovery rate (FDR) [21].

All data processing was performed with MATLAB (version 9.0, The Mathworks, Inc., Natick, Massachusetts, United States)

3.2.3 Ethical statement

All data used in this study was anonymized prior to their uploading to the 1000 functional connectome database. No identifying data was requested, and all data used is publicly available through the 1000 functional connectome website[1]. This study was approved by the Human Investigation Committee of Yale University as a 'not human subjects research'.

3.3 Results

Regression methods are inherently designed to remove part of the signal. To get a rough estimate of the amount of signal removed by a regression, the time-courses of a single, representative subject (New Haven dataset) were brought to a similar scale by setting the mean to 1 to subsequently represent the spontaneous BOLD signal fluctuations as percentage change (Figure 2). After all regressions a large drop in signal intensity occurs, indicating that functional connectivity (after regression) is based on a relatively very small part of the fMRI signal. For example, with global signal

[1] See: https://www.nitrc.org/projects/fcon_1000/

regression BOLD signal amplitude in gray matter drops by a factor of 8, whereas with white matter signal regression BOLD signal amplitude drops by a factor of 2.

As a measure of data quality, the root-mean-square (RMS) of the percentage change in raw global signal was used. Bangor and Munich have a far greater RMS (Figure 3, Figure 4), indicating far larger variability in their global signals. These results imply that there are subsets of data that could have potential artefacts in the Bangor and Munich datasets. The origin of these artefacts is unknown and they could not be resolved thus the Bangor and Munich datasets were excluded from all statistical analyses.

Whilst regressions remove a large part of the signal, the resulting correlation maps are similar across the experiment sites included in statistical analyses (Figure 5). All regressions reduce correlations across the brain, generate correlation maps with similar spatial distributions, and reduce correlations within the DMN and TPN with their respective averaged time-course (Figure 6). One-tailed t-tests of the differences in z-scores were done in both directions for each network and each site. Resulting p-values in the same direction were averaged using Fisher's method and the combined p-values were used to test for significance ($p < .05$, FDR corrected, Bangor and Munich excluded) Whilst there are significant differences between regressions in both the DMN and TPN (DMN: raw > WMSR > GMSR > GSR; TPN: raw > WMSR = GMSR > GSR), these differences are marginal compared to the difference between no regression and any other regression (Table 2A, Table 2B).

Global signal regression supposedly reduces the effect of motion on the BOLD signal [14]. To test whether other signal regressions have a similar effect, the correlation between motion parameters and all voxels' time-series were calculated (Figure 7). One-tailed two-sample t-tests were performed for all regression pairs in each experimental site and in both directions. P-values across experimental sites in the same direction were averaged with Fisher's method. All regressions significantly reduced the averaged correlation of voxel time-series with all motion parameters, and statistical difference between regressions exist (Raw > GMSR > GSR > WMSR, $p < .05$, FDR corrected, Bangor and Munich

excluded). However, the differences between regressions are small compared to the difference between no regression and any regression (Table 2C).

3.4 Discussion

Region specific regressions are a commonly used step in R-fMRI signal preprocessing. Here, the effects of these regressions on functional connectivity and motion correction were analyzed. GMSR, GSR, and WMSR all reduce correlations of DMN and TPN voxels with their respective averaged network signals and they reduce correlations of all voxels with the motion parameters. Furthermore, these regressions produce similar spatial correlation patterns. These results indicate that GMSR, GSR, and WMSR produce comparable effects. This work suggests that, in studies of functional connectivity under disease, any of these regressions can be used to produce similar results, thus avoiding regions of concern to the disease state.

A difference could exist between the effects of region specific regressions on certain signal frequencies. If this were true, then filtering the data could interact with the regressions and thus introduce effects that were not observed here. For example, if the regressions in the Munich dataset primarily affected frequencies above 0.1 Hz, then a filtering step might make the Munich dataset more similar to the included datasets. Data filtering was not done here to prevent interaction effects between filtering and regression, and these interaction effects are left for future research.

GSR reduces some metabolic components of the BOLD signal [12], it is correlated with local field potentials [13], and reduces the effect of motion on the BOLD signal [14]. Given the similarity between the raw global signal, the raw gray matter signal, and the raw white matter signal and the similarity of GSR, GMSR and WMSR in our results, it is plausible that the properties of GSR hold for WMSR and GMSR as well. Indeed, a reduction of the effect of motion was found after all regressions, and WMSR had a similar effect on the metabolic components of the BOLD signal [12].

Region specific signal regressions generally have a similar effect within datasets, but can have very different effects on different datasets. The cause of this lies in large signal fluctuations only present in the deviating datasets. These fluctuations may be due to bad shim, motion artefacts, a flaw in the scanner, or another unknown artefact. However, without all details of the scans finding the cause of these fluctuations is challenging. Despite the different field strengths, gender ratios, subject ages, and (likely) demographics in these datasets, region specific signal regressions produce remarkably similar correlation patterns across these datasets, indicating that any of the tested regressions can be used to produce similar results regardless of scanning parameters. These results indicate that R-fMRI studies using a regression can be compared across experimental sites, regardless of those sites' scanning parameters.

4 METABOLIC STATE ALTERS THE COST OF FUNCTIONAL CONNECTIVITY

4.1 Introduction

R-fMRI is a neuroimaging method used to analyze spontaneous fluctuations in the BOLD signal. Correlations of the time-series of every voxel with a signal of interest can be used to find brain networks [22], or to analyze interactions between networks [10]. Whilst spontaneous BOLD signal fluctuations are widely used, their exact origin is unclear.

The amplitude of low frequency fluctuations (ALFF) is a proposed measure of regional spontaneous activity in R-fMRI data [23, 24]. ALFF is positively correlated with CMR_{glc} within subjects [9], and subjects in a more active metabolic state have higher ALFF than those in a less active metabolic state [12].

fMRI data has to undergo several preprocessing steps before it can be analyzed. These steps typically consist of, but are not limited to, blurring, regression, filtering, and normalization. Whilst these preprocessing steps are used extensively, no one has systematically analyzed the effects of these preprocessing steps on functional connectivity. In this chapter the effects of signal regression and metabolic state on the correlation between ALFF and CMR_{glu} will be analyzed. The filtering step was not analyzed because the ALFF algorithm has an in-built filter. The normalization step standardizes signal intensity across voxels and thus forces ALFF to be approximately equal in all voxels, hence normalization was not analyzed either in this work.

4.2 Methods

4.2.1 Data Collection

A summary of the human data collection is provided here. For a more detailed description of data acquisition see this dataset's first publication [8].

Twenty-four healthy subjects were recruited of which two were excluded. Fludeoxyglucose positron emission tomography (FDG-PET), anatomical MRI and R-fMRI were simultaneously acquired with an integrated Siemens Biograph mMR scanner. Subjects were split into two equally sized groups. One group was scanned with eyes closed (8 male, 3 female, mean±s.d.: 52.2±10.4 years old, 77.0±13.3 kg, 177.7±9.1 cm tall), the other group was scanned with eyes open (7 male, 4 female, mean±s.d.: 56.7±9.6 years old, 75.4±16.4 kg, 172.7±7.8 cm tall). Subjects were instructed to keep their eyes closed/open, relax, and not fall asleep. A T2*-weighted resting state scan was performed (EPI, TR 2s, TE 30ms, 35 slices with 0.6mm gap, 192x192mm FOV, 64x64 matrix size, 300 volumes, 10 minutes 8 seconds). A bolus injection of FDG tracer was given at the start of the resting state scans. Following the resting state scan, subjects were instructed to close their eyes and a T1-weighted anatomical scan was acquired (MP-RAGE, TR 2.3s, TE 2.98ms, 160 slices with 0.5mm gap, 256x256mm FOV, 256x256 matrix size, 5 minutes 3 seconds). Thirty minutes after the bolus injection, FDG-PET images were acquired (saturated list mode, 128 slices with 0.5mm gap, 192x192mm matrix size, 3.7x2.3x2.7 mm3 voxel sizes, 10 minutes).

Scans were performed with all lights off. Before the resting state scan the light from the console room was used for the bolus injection.

4.2.2 Preprocessing and network creation

Registration and network creation were performed as described by [12]. In short, R-fMRI data were slice-timing and motion corrected. Individual subjects' data were segmented to produce masks of gray matter, white matter, cerebrospinal fluid, and the whole-brain. R-fMRI and FDG-PET data were registered to each subject's gray matter mask. The data in anatomical space was registered to the MNI template (2mm, isotropic). Registered data was blurred with a Gaussian filter (full width half maximum 8mm, size 6mm). Two separate regressions were run: a regression with the six motion parameters and a regression of the six motion parameters, the global signal, white matter signal and cerebral spinal fluid signal.

Blurred (not regressed or filtered) R-fMRI data in MNI space and the Group ICA of fMRI toolbox (GIFT, mialab.mrn.org/software/gift/) [20] were used to generate 20 independent component (IC) networks (for network classification see [12]). Networks consisting primarily of thermal noise, white matter or CSF were removed.

4.2.3 Data Analysis

BOLD amplitude is calculated similarly to the amplitude of low frequency fluctuations (ALFF) [23, 24], with the exceptions that no normalization is performed, and a frequency pass-band between .01 and .06 Hz was used. This pass-band was chosen to match the pass-band used by [9]. To calculate BOLD amplitude, the time-series of each voxel is transformed to the frequency domain with a fast Fourier transformation. BOLD amplitude is the mean square root of all frequencies within the pass-band.

Before calculation of BOLD amplitude, all time-series were divided by the mean of all time-series across all brain voxels, time points and subjects of the same species[2]. This was done to make across species BOLD amplitude more comparable. Power in the frequency domain scales linearly with amplitude in the time domain, so the scaling across species does not affect BOLD amplitude ratios within species and thus does not affect any of the analyses.

FDG-PET data were left in the original units, Becquerel per milliliter (Bq/ml). For a comparison between Bq/ml and quantitative units, see [12].

Best-fit lines were calculated with a geometric mean regression [25], also known as reduced major axis regression. Contrary to least-squares regression, geometric mean regression does not assume that the x-axis is measured without error, and the best-fit line is independent of choice of axes, i.e. when the axes are flipped, the best fit line remains the same [25]. Both BOLD amplitude and FDG are measured with error and there is no reason to prefer either as the x-axis thus the geometric mean

[2] The research presented here is part of a larger research involving, amongst others, resting state data of rats. As no part of the rat research was conceived, gathered, analyzed, or written up by me, the rat data is not shown here.

regression provides a more appropriate best fit line than least-squares regression [26].

Multiple comparison correction was done with the sequential goodness of fit algorithm (SGoF) [27]. SGoF is a metatest of p-values that increases its power with the number of tests. The SGoF algorithm uses a binomial or chi-square test to estimate whether the number of significant results differs significantly from the expected number of significant results. If there is a significant difference between these, then the p-value threshold is modified and the algorithm runs again until the observed and expected number of significant results no longer differ significantly. SGoF was chosen over FDR as it performs better with smaller effect sizes [27].

All data processing was performed with MATLAB (version 9.0, The Mathworks, Inc., Natick, Massachusetts, United States), except for the MANOVA which was performed with R [28].

4.3 Results

BOLD amplitude and FDG were plotted for both eyes open and eyes closed after motion parameter regression (Figure 8a). Both BOLD amplitude and FDG are highest in cortical areas. BOLD amplitude and FDG within the gray matter of each ICA network, except for #1, #3, #6, #20 (see Methods for exclusion criteria), are plotted (Figure 8b). Pearson correlations between BOLD amplitude and FDG were r=0.42 and r=0.59 for eyes closed and open, respectively, indicating that a linear relation exists between them. A MANOVA revealed a significant main effect of both metabolic state and ROI ($p < 0.05$, Pillai's trace, SGOF-corrected), but no interaction effect ($p > 0.05$, Pillai's trace, SGOF-corrected). Regression of the global, white matter, and cerebral spinal fluid signals, and the motion parameters yielded Pearson correlations of r=0.41 and r=0.54 for eyes closed and open, respectively. The correlations without regression were r=0.39 and r=0.56 for eyes closed and open, respectively. Together, these correlations indicate that regression has little effect on how well BOLD amplitude represents metabolic activity.

4.4 Discussion

Consistent with previous results, a linear relationship between BOLD amplitude and FDG was found [9], and an increase of both BOLD amplitude and FDG was observed in a more active metabolic state [12] ([12] used the same dataset). This work is the first to show that a linear relationship holds across different metabolic states, even though the coupling between BOLD amplitude and FDG is different in these metabolic states.

Surprisingly, signal regression had little effect on the correlation between BOLD amplitude and FDG within condition. Results from the same dataset have previously shown that signal regression removed the statistically significant difference between ALFF in the eyes open and closed conditions [12]. One possible cause of this is that signal regression primarily causes a mean shift in the frequency bands included in ALFF, but doesn't meaningfully change their distribution. As Pearson correlations do not change with mean shifts, this could leave the correlations intact whilst simultaneously decreasing the difference between metabolic states.

Together, these results show that BOLD amplitude is an indicator of metabolic activity, regardless of regression method or metabolic state.

5 CONCLUSION

In this thesis, the effect of different types of signal regression on R-fMRI data were analyzed and the effect of motion parameters regression with and without region specific regression on the coupling between BOLD amplitude and metabolic activity was considered. Whilst there was a large difference between any regression and no regression, the type of region specific regression had little effect on the resulting correlation maps and within-network correlations. Also, region specific regressions may hide artefacts in the raw data, thus calling for careful analysis of the raw data before publishing resting-state data. Lastly, regressions had little effect on the correlation between BOLD amplitude and metabolic activity, indicating that some metrics may be largely unaffected by region-specific regressions.

Together, these projects show that: 1) in studies of functional connectivity under disease, any of the studied regressions may be used, thus avoiding regions of concern to the disease state, 2) any regression will increase anti-correlations in the data, and 3) regardless of the mode of regression, BOLD amplitude is an indicator of metabolic activity, but the linear coupling is dependent on the subjects' metabolic state.

6 ACKNOWLEDGEMENTS

First and foremost, I would like to thank my direct supervisors prof. Fahmeed Hyder, and dr. Garth Thompson for their amazing guidance, and my local supervisor Remco Renken for his valuable input. I also thank the rest of the Hyder lab for their support, and prof. Robert Fulbright for assisting with the forms for the Human Investigation Committee.

7 REFERENCES

1. Kennedy, C., et al., *Mapping of functional neural pathways by autoradiographic survey of local metabolic rate with (14C) deoxyglucose.* Science, 1975. **187**(4179): p. 850-853.
2. Kadekaro, M., A.M. Crane, and L. Sokoloff, *Differential effects of electrical stimulation of sciatic nerve on metabolic activity in spinal cord and dorsal root ganglion in the rat.* Proceedings of the National Academy of Sciences, 1985. **82**(17): p. 6010-6013.
3. Hyder, F., et al., *Neuronal–glial glucose oxidation and glutamatergic–GABAergic function.* Journal of Cerebral Blood Flow & Metabolism, 2006. **26**(7): p. 865-877.
4. Hyder, F., D.L. Rothman, and M.R. Bennett, *Cortical energy demands of signaling and nonsignaling components in brain are conserved across mammalian species and activity levels.* Proceedings of the National Academy of Sciences, 2013. **110**(9): p. 3549-3554.
5. Hyder, F., et al., *Uniform distributions of glucose oxidation and oxygen extraction in gray matter of normal human brain: No evidence of regional differences of aerobic glycolysis.* Journal of Cerebral Blood Flow & Metabolism, 2016: p. 0271678X15625349.
6. Sokoloff, L., *Energy Metabolism in Neural Tissues in vivo at Rest and in Functionally.* Brain Energetics and Neuronal Activity: Applications to fMRI and Medicine, 2004: p. 11.
7. Shu, C.Y., et al., *Quantitative β mapping for calibrated fMRI.* NeuroImage, 2016. **126**: p. 219-228.
8. Riedl, V., et al., *Local activity determines functional connectivity in the resting human brain: a simultaneous FDG-PET/fMRI study.* J Neurosci, 2014. **34**(18): p. 6260-6.
9. Tomasi, D., G.-J. Wang, and N.D. Volkow, *Energetic cost of brain functional connectivity.* Proceedings of the National Academy of Sciences, 2013. **110**(33): p. 13642-13647.
10. Fox, M.D., et al., *The human brain is intrinsically organized into dynamic, anticorrelated functional networks.* Proceedings of the National Academy of Sciences of the United States of America, 2005. **102**(27): p. 9673-9678.
11. Murphy, K., et al., *The impact of global signal regression on resting state correlations: are anti-correlated networks introduced?* Neuroimage, 2009. **44**(3): p. 893-905.
12. Thompson, G.J., et al., *The whole-brain "global" signal from resting state fMRI as a potential biomarker of quantitative state changes in glucose metabolism.* Brain Connectivity, 2016(ja).
13. Schölvinck, M.L., et al., *Neural basis of global resting-state fMRI activity.* Proceedings of the National Academy of Sciences, 2010. **107**(22): p. 10238-10243.
14. Power, J.D., et al., *Methods to detect, characterize, and remove motion artifact in resting state fMRI.* Neuroimage, 2014. **84**: p. 320-341.
15. Patriat, R., et al., *The effect of resting condition on resting-state fMRI reliability and consistency: a comparison between resting with eyes open, closed, and fixated.* Neuroimage, 2013. **78**: p. 463-473.
16. Brown, V.M., et al., *Altered resting-state functional connectivity of basolateral and centromedial amygdala complexes in posttraumatic stress disorder.* Neuropsychopharmacology, 2014. **39**(2): p. 361-369.
17. Gotts, S.J., et al., *Fractionation of social brain circuits in autism spectrum disorders.* Brain, 2012. **135**(9): p. 2711-2725.
18. Satterthwaite, T.D., et al., *Heterogeneous impact of motion on fundamental patterns of developmental changes in functional connectivity during youth.* Neuroimage, 2013. **83**: p. 45-57.
19. Biswal, B.B., et al., *Toward discovery science of human brain function.* Proceedings of the National Academy of Sciences, 2010. **107**(10): p. 4734-4739.
20. Correa, N., T. Adalı, and V.D. Calhoun, *Performance of blind source separation algorithms for fMRI analysis using a group ICA method.* Magnetic resonance imaging, 2007. **25**(5): p. 684-694.

21. Benjamini, Y. and Y. Hochberg, *Controlling the false discovery rate: a practical and powerful approach to multiple testing.* Journal of the Royal Statistical Society. Series B (Methodological), 1995: p. 289-300.
22. Biswal, B., et al., *Functional connectivity in the motor cortex of resting human brain using echo‐planar mri.* Magnetic resonance in medicine, 1995. **34**(4): p. 537-541.
23. Zang, Y.-F., et al., *Altered baseline brain activity in children with ADHD revealed by resting-state functional MRI.* Brain and Development, 2007. **29**(2): p. 83-91.
24. Zou, Q.-H., et al., *An improved approach to detection of amplitude of low-frequency fluctuation (ALFF) for resting-state fMRI: fractional ALFF.* Journal of neuroscience methods, 2008. **172**(1): p. 137-141.
25. Halfon, E., *Regression method in ecotoxicology: A better formulation using the geometric mean functional regression.* Environmental science & technology, 1985. **19**(8): p. 747-749.
26. Smith, R.J., *Use and misuse of the reduced major axis for line‐fitting.* American Journal of Physical Anthropology, 2009. **140**(3): p. 476-486.
27. Carvajal-Rodríguez, A., J. de Uña-Alvarez, and E. Rolán-Alvarez, *A new multitest correction (SGoF) that increases its statistical power when increasing the number of tests.* BMC bioinformatics, 2009. **10**(1): p. 1.
28. R Core Team, *R: A language and environment for statistical computing*. 2016, R Foundation for Statistical Computing: Vienna, Austria.

8 FIGURES

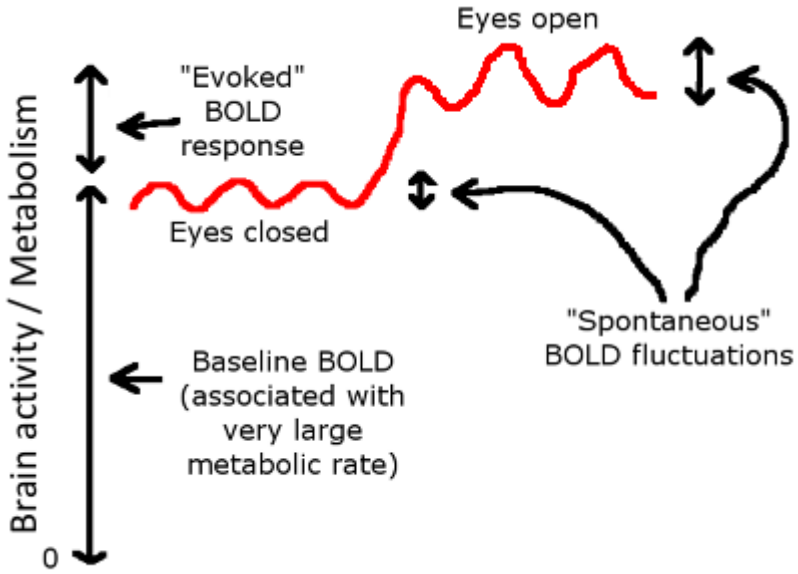

Figure 1. Schematic representation of the BOLD signal in different metabolic sates. During the eyes closed condition the baseline BOLD signal is lower and has smaller spontaneous fluctuations than in the eyes open condition. This indicates that the baseline of the signal and its fluctuations are somehow linked. This figure is a modified version of a figure provided by Garth Thompson. The only modification was an incorporation of the increased amplitude of spontaneous BOLD fluctuations in the eyes open condition.

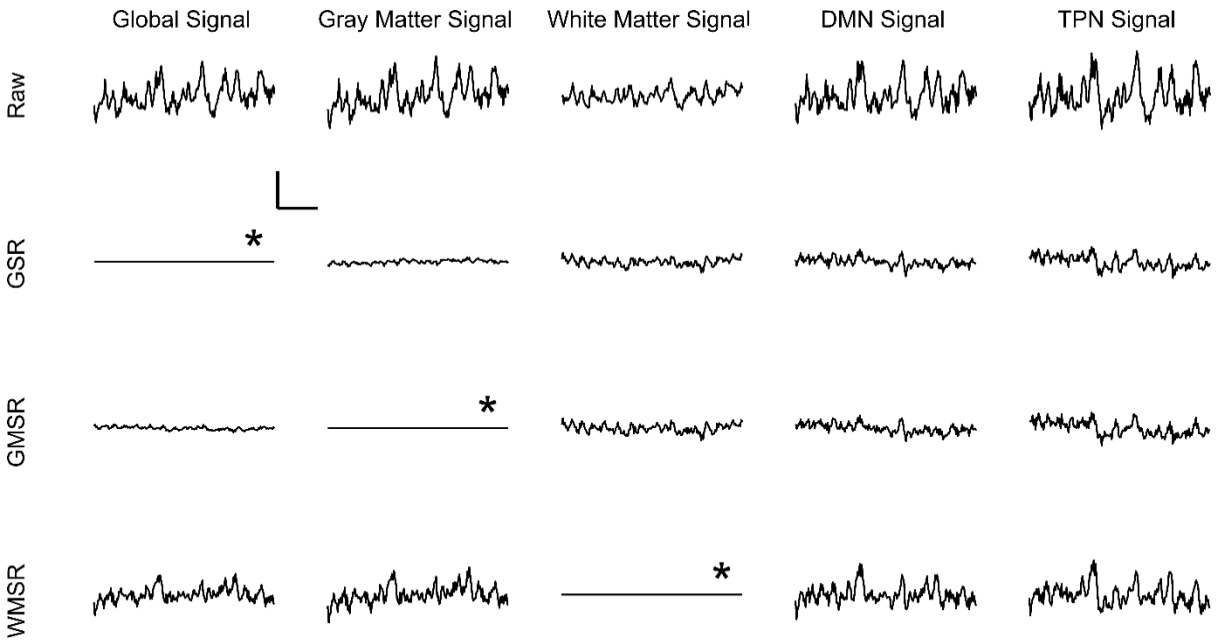

Figure 2. Effects of region-specific regression on BOLD signal amplitude. Data from a single subject (New Haven dataset). Averaged BOLD signal time-courses in the (A) whole brain, (B) gray matter, (C) white matter, (D) default-mode network (DMN), and (E) task-positive network (TPN). The BOLD signal time courses shown are raw data before regression (first row), after global signal regression (GSR, second row), after gray matter signal regression (GMSR, third row), and after white matter signal regression (WMSR, fourth row). The horizontal bar denotes 60 seconds and the vertical bar denotes 300 (arbitrary units). Asterisks denote that these signals are expected to be flat because these were the regional signals that were regressed.

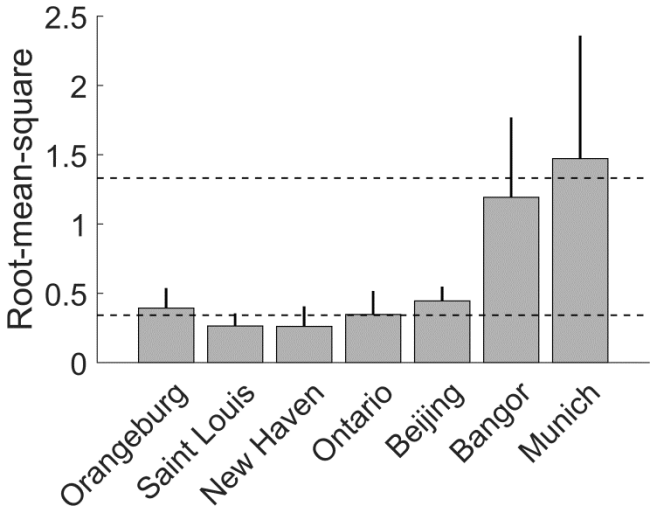

Figure 3. Assessment of BOLD signal variability. Root-mean square (RMS) of raw global BOLD signals expressed as percentage signal change from the mean. Error bars represent standard deviation across all subjects in each data set. Most of the datasets (Orangeburg, Saint Louis, New Haven, Ontario, Beijing) showed BOLD signal amplitudes that were ~0.5%, which is within general observations from various R-fMRI studies. However other datasets (Bangor, Munich) showed BOLD signal amplitudes that were significantly larger (>1%). Closer inspection of the Bangor and Munich datasets show that some subjects had very large fluctuations in BOLD signal that are beyond the level expected within physiological limits (Figure 4). The top dashed line is the mean of the Bangor and Munich datasets, whereas the bottom dashed line is the mean of all other datasets. RMS of the Bangor and Munich datasets was much statistically much higher, indicating larger variability in their global BOLD signal amplitude. Experiment sites left of the vertical line are sorted by intensity of the raw image.

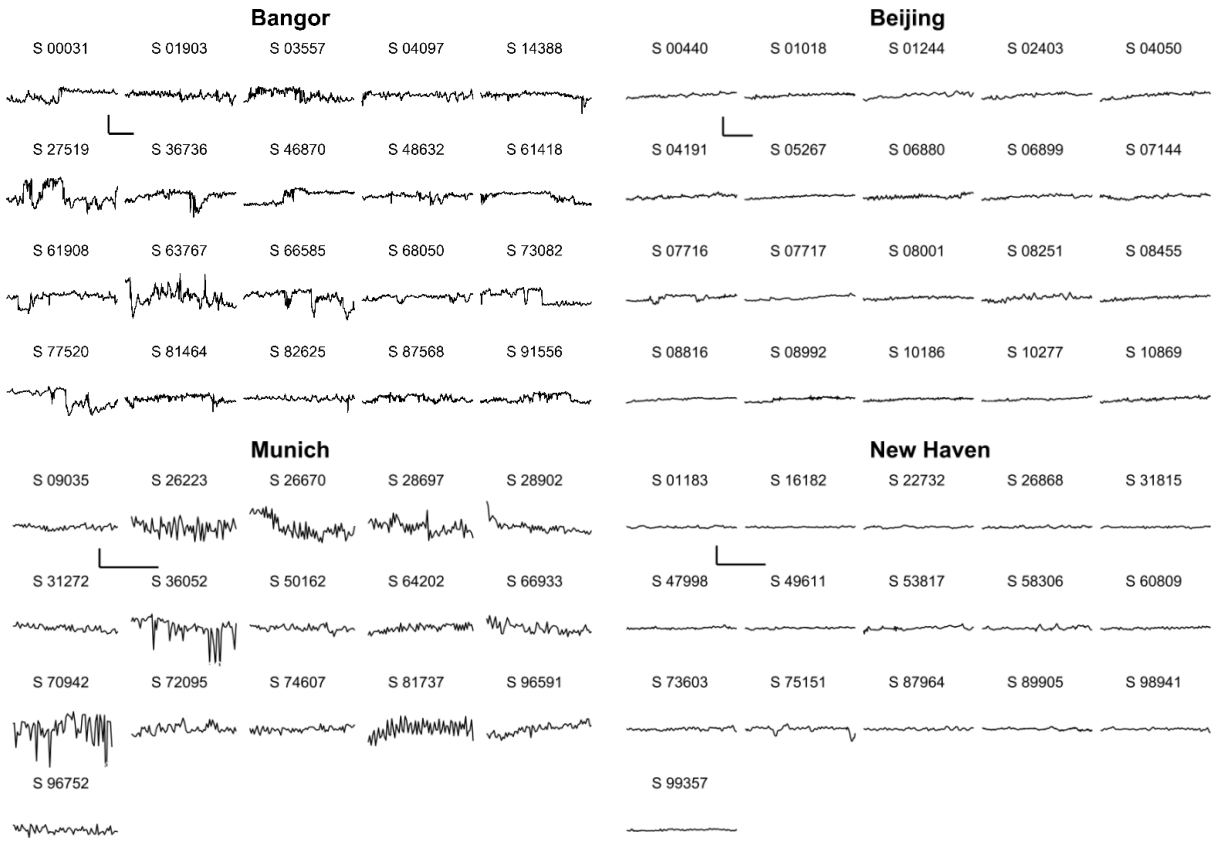

Figure 4. Comparison of raw global BOLD signal fluctuations per subject across specific data sets. Top row shows Bangor and Beijing datasets, which have 20 subjects each. Bottom row shows Munich and New Haven datasets, which have 16 subjects each. Some subjects in the Bangor and Munich datasets have BOLD signal fluctuations that are physiologically implausible, indicating some unknown artefact in these data. By comparison, all subjects in the Beijing and New Haven datasets are relatively stable during their entire scan for each subject. S-numbers refer to the subject numbers as documented by the 1000 functional connectomes project. The vertical bars represent 5% signal change, whereas the horizontal bars represent 120 seconds for their respective datasets.

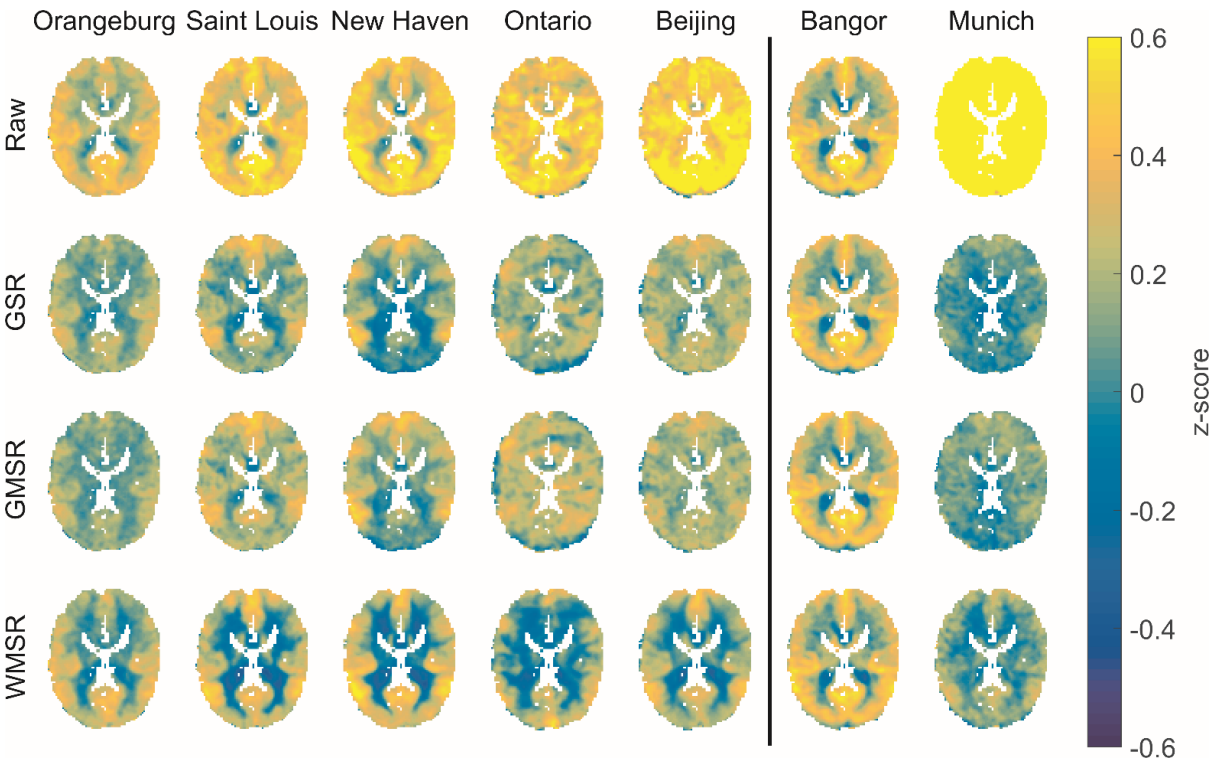

Figure 5. Effects of region-specific regression on appearance of negative functional connectivity. Pearson correlations (expressed in z-score) of all voxels' time-series with the default mode network (DMN) averaged signal are shown for all experimental sites (columns) in the raw data (first row), and after regression of the global signal (GSR, second row), the gray matter signal (GMSR, third row), and the white matter signal (WMSR, fourth row). All z-scores are averaged across subjects for each dataset. Signal regressions reveal comparable correlation patterns across the five datasets left of the vertical line. The Bangor dataset barely changes with any regression, and the Munich dataset has high correlations throughout the entire brain that are removed with any signal regression. By itself, any regressed slice would pass a visual inspection, but when all regressions are shown the deviant datasets are easily identified. These observations shown for one slice are typical for all slices in the brain.

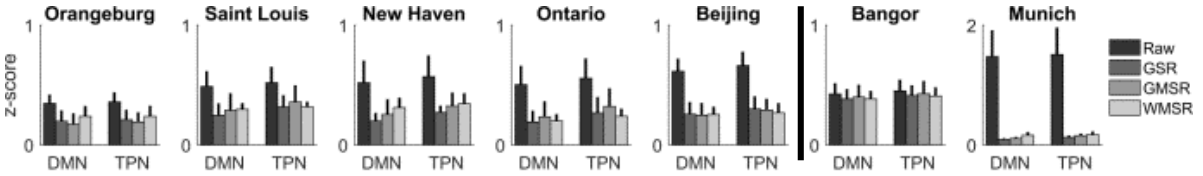

Figure 6. Change in voxels' correlation with their own network after regression. Pearson correlations (expressed in z-score) of voxels in the default mode network (DMN) and the task positive network (TPN) with their respective averaged network signal are shown. Data is shown for the raw signal, and for the signals after regression of the global signal (GSR), the gray matter signal (GMSR), and the white matter signal (WMSR). Error bars represent standard deviation across subjects. In the five left-most datasets any regression reduces correlations by approximately a factor 2. Marginal statistically significant differences exist (DMN: raw > WMSR > GMSR > GSR; TPN: raw > WMSR = GMSR > GSR, Bangor and Munich excluded in statistical analyses). Similar to previous results in Figure 5, regressions have little effect on the Bangor dataset and a huge effect on the Munich dataset, giving further indication that these datasets are outliers.

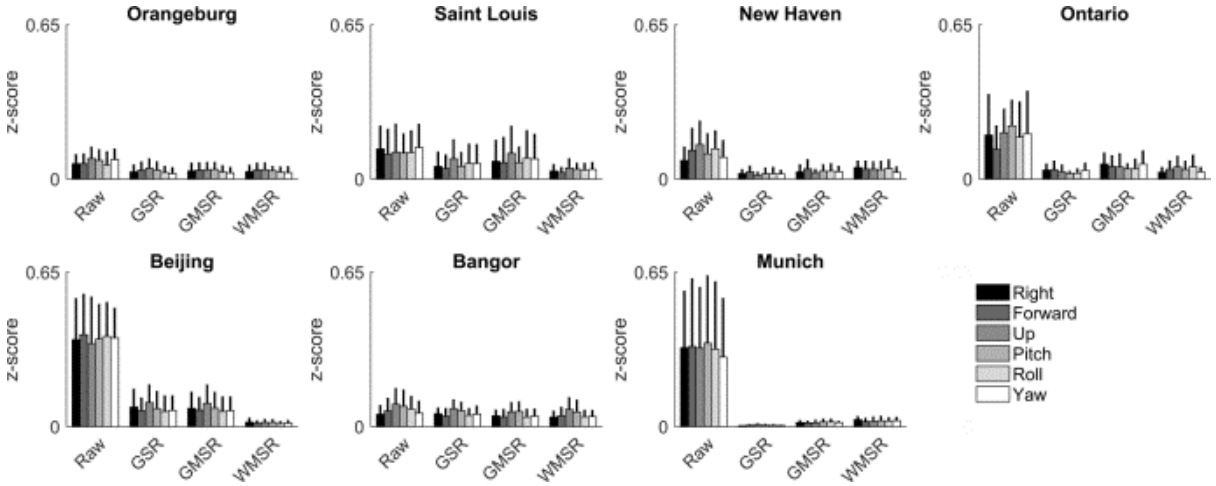

Figure 7. Regressions correct for motion artefacts. Absolute Pearson correlations (expressed in z-score) of all voxels' time-series with rigid-body motion parameters were averaged across subjects. Error bars represent standard deviations across subjects. Correlations between motion parameters and voxels' time-series in the raw data are consistently larger than those after global signal regression (GSR), gray matter signal regression (GMSR), and white matter signal regression (WMSR). Statistically significant differences between regressions exist (GMSR > GSR > WMSR), but are marginal compared to the difference between no regression and any other regression, indicating that, for purposes of motion correction, these regressions produce comparable results. Similar to prior results in Figure 5 and Figure 6, signal regressions have little effect on the Bangor dataset and a large effect on the Munich dataset.

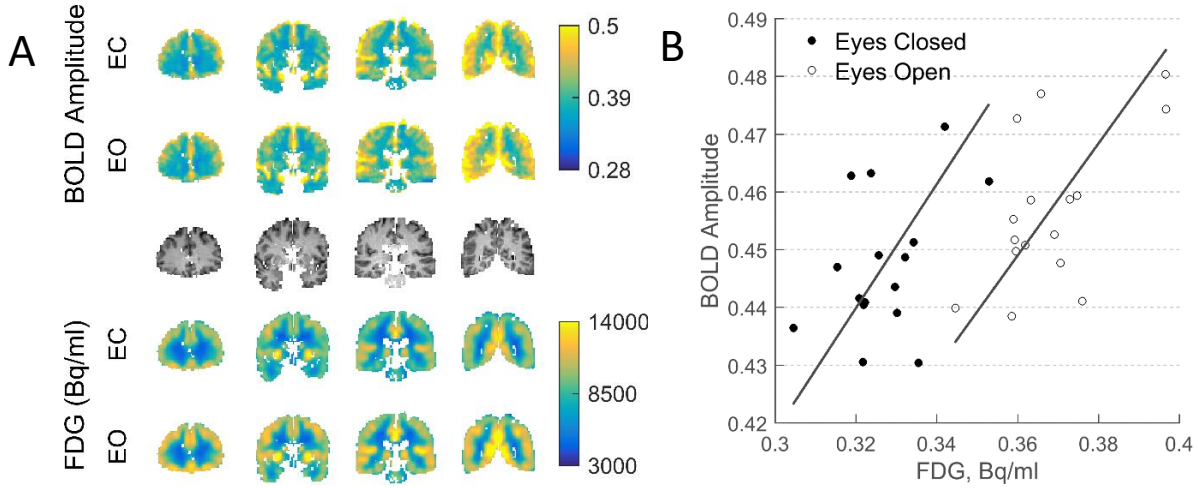

Figure 8. Relationship between BOLD amplitude and FDG. (A) Coronal slices (slices 30, 45, 60, 75) in 2mm MNI space of BOLD amplitude after motion parameter regression (top two rows), average anatomical images (middle row), and FDG (bottom two rows) in the eyes closed (EC) and eyes open (EO) conditions. ALFF and FDG have comparable spatial distribution. (B) Scatterplot of FDG vs BOLD amplitude after motion parameter regression in ICA networks averaged over subjects. Best fit lines were calculated with geometric mean regression. A linear relationship is observed between BOLD amplitudes and metabolic activity. In addition, both modalities slightly, but significantly, increased in the EO state.

9 TABLES

Table 1. Details of all used 1000 Functional Connectome Project datasets. A "\"denotes missing data. Sample sizes were calculated after subject exclusions.

Dataset	Magnet	N	Sex	Age Range	TR	Time-points	eyes open vs. closed	slice acquisition order	handedness
Bangor, UK	3T	20	0F	19-38	2	265	open	sequential ascending	Right-handed
Beijing (Zang), China	3T	20	11F	18-26	2	225	closed	interleaved ascending	Right-handed
Munich, Germany	1.5T	16	6F	63-73	3	72	closed	interleaved ascending	Right-handed
New Haven (B), USA	3T	16	8F	18-42	1.5	181	open, no projection	interleaved ascending	Right-handed
Saint Louis, USA	3T	16	16F	21-29	2.5	127	open, fixation	interleaved ascending	Right-handed
Ontario, Canada	4T	9	\	\	3	105	closed	segmented spiral, 2-shot interleaved, descending	Mixed
Orangeburg, USA	1.5T	19	5F	20-55	2	165	closed	Interleaved ascending	Mixed

Table 2. Effect of signal regression on (A) network coherence in the default mode network, (B) network coherence in the task positive network, and (C) correlation with the motion parameters. The top half of both tables shows the change in average correlation of all voxels within each network with the averaged network signal. All percentages are expressed as row / column. The bottom half of both tables contains the lowest p-value of the one-tailed two-sample t-tests between regressions. The differences between the raw correlations and any of the regressed correlations are far larger than any other difference. Asterisks denote significant p-values after multiple comparison correction.

A

Default-mode network

	Raw	GSR	GMSR	WMSR
Raw		227%	207%	194%
GSR	≈ 0*		93%	85%
GMSR	≈ 0*	.0002*		93%
WMSR	≈ 0*	≈ 0*	.0010*	

B

Task positive network

	Raw	GSR	GMSR	WMSR
Raw		194%	181%	191%
GSR	≈ 0*		94%	98%
GMSR	≈ 0*	≈ 0*		106%
WMSR	≈ 0*	.0014*	.0985	

C

Motion Correction

	Raw	GSR	GMSR	WMSR
Raw		27%	33%	20%
GSR	≈ 0*		123%	75%
GMSR	≈ 0*	0.002*		61%
WMSR	≈ 0*	0.005*	≈ 0*	