Effective Connectivity in Depression

Supplementary Information

Supplementary Methods

Participants

There were 336 patients with a diagnosis of major depression, and 350 controls. The patients were from Xinan (First Affiliated Hospital of Chongqing Medical School in Chongqing, China), and Taiwan (Veteran General Hospital, Taipei). All participants were diagnosed according to the Diagnostic and Statistical Manual of Mental Disorder-IV criteria for major depressive disorder. Depression severity and symptomatology were evaluated by the Hamilton Depression Rating Scale (HAMD, 17 items) (1) and the Beck Depression Inventory (BDI) (2). Table S1 provides a summary of the demographic information and the psychiatric diagnosis (showing how they were diagnosed) of the participants. The data collection was approved by the local ethical review committees, was in accordance with the Code of Ethics of the World Medical Association (Declaration of Helsinki), and informed consent was obtained. This is a subset of patients from our previous functional connectivity investigation (3), but the analysis used here is completely different and novel in its application to depression, for it assesses effective or directed connectivity, and goes beyond correlations. With respect to age and sex, Table S1 shows that there were no significant differences in the age and sex of the depressed groups and the controls. Further, the effects of age and sex were regressed out in all analyses and the comparison of results between male and female, younger and older participants are shown in Fig. S5 and Fig. S6. 125 of the patients were not receiving medication at the time of the neuroimaging. Further details follow.

Xinan: Patients with MDD were recruited from the outpatient department of the First Affiliated Hospital of Chongqing Medical School in Chongqing, China. All were diagnosed according to the Structured Clinical Interview for DSM-IV, by independent assessments of two psychiatrists. They were also assessed for disease severity using the Hamilton Depression Rating Scale

(HAMD) (1) and Beck Depression Inventory (BDI), illness duration and the medication status of the patients. Before the investigation, we excluded individuals who were not suitable for MRI scanning by interview and by the self-reported checklist. The MRI related exclusion criteria include claustrophobia, metallic implants, Meniere's Syndrome and a history of fainting within the previous half year. Exclusion criteria for both groups were as follows: current psychiatric disorders (except for MDD) and neurological disorders; substance abuse; and stroke or serious encephalopathy. Of note, all of the subjects in the control group did not meet DSM-IV criteria for any psychiatric disorders and did not use any drugs that could affect brain function. This study was approved by the Research Ethics Committee of the Brain Imaging Center of Southwest University and First Affiliated Hospital of Chongqing Medical School. Informed written consent was obtained from each subject. This study was conducted in accordance with the Helsinki Declaration as revised in 1989.

Taiwan: Patients were recruited from the Veteran General Hospital in Taipei, Taiwan. All participants were diagnosed according to the Diagnostic and Statistical Manual of Mental Disorder-IV criteria for depression, and each participant's history of medical disease, psychiatric illness, and medication use was evaluated by interview and medical charts carefully. Experiments were conducted in accordance with the Declaration of Helsinki and approved by the Institutional Review Board of Taipei Veterans General Hospital. Written informed consent was obtained from all participants after ensuring adequate understanding of the study. Any participants with the following conditions were excluded: 1) a comorbid substance-related disorder, 2) presence of neurobiological disorders, such as dementia, head injury, stroke, or Parkinson's disease; 3) presence of hypertension, diabetes, hyperlipidemia or coronary heart disease; 4) severe medical illness, such as malignancy, heart failure, or renal failure; 4) presence of ferromagnetic foreign bodies or implants that were anywhere in the body. Depression severity was evaluated by the psychiatrist-assessed Hamilton Depression Rating Scale (HAMD, 17 items) (1).

Table S1. A summary of the demographic information and the psychiatric diagnosis in the present study.

Sites	Group	Age (years)	Sex (male/female)	Education (years)	Medication (yes / no)	HAMD	BDI	Duration of illness	Mean FD
Taiwan	Healthy	49.18±8.58	60 / 36	15.04 ± 3.83	/	/	/	/	0.133 ± 0.054
	Patient	52.64±14.86	33 / 21	12.66 ± 3.95	54 / 0	9.34 ± 6.99	/	8.63 ± 6.92	0.116 ±0.056
	Statistic (t / p)	-1.810 / 0.072	0.028 / 0.866	3.60 / 4.3e-4		/	/	/	1.833 / 0.0687
Xinan	Healthy	39.65 ± 15.80	166 / 88	13.01 ± 3.89	/	/	/	/	0.133 ±0.063
	Patient	38.74 ± 13.65	183 / 99	11.91 ± 3.58	157 / 125	20.8 ± 5.87	20.42 ± 9.33	4.16 ± 5.51	0.125 ± 0.054
	Statistic (t / p)	0.719 / 0.472	0.013 / 0.911	3.41 / 6.9e-4	/	/	/	/	1.729 / 0.084

Image Acquisition and Preprocessing

Data for resting state connectivity analysis were collected in 3T MRI scanners in an 8 min period in which the participants were awake in the scanner not performing a task using standard protocols described here.

Data preprocessing was performed using DPARSF (4) (http:// restfmri.net) which is a toolbox based on the SPM8 software package. The first 10 EPI scans were discarded to suppress equilibration effects. The remaining scans of each subject underwent slice timing correction by sinc interpolating volume slices, motion correction for volume to volume displacement, spatial normalization to standard Montreal Neurological Institute (MNI) space using affine transformation and nonlinear deformation with a voxel size of $3 \times 3 \times 3mm^3$, followed by spatial smoothing (8 mm Full Width Half Maximum). To remove the sources of spurious correlations present in resting-state BOLD data, all fMRI time-series underwent band-pass temporal filtering (0.01-0.1 Hz), nuisance signal removal from the ventricles, and deep white matter, and regressing out any effects of head motion using the Friston et al 24 head motion parameters procedure (5). We note that with a TR of 2 s the effective connectivity method may benefit from band-pass temporal filtering in the range 0.01-1.0 Hz, to increase the use of the high frequency fluctuations in the times series, but with the filtering applied did obtain sufficient information to obtain asymmetric weights in the links between connected brain areas to be detected. This filtering used though did preserve sufficient of the high-frequency fluctuations that can be successfully captured by the model.

Finally, we implemented additional careful volume censoring ("scrubbing") movement correction as reported by Power et al. (6) to ensure that head-motion artifacts are not driving observed effects. The mean framewise displacement (FD) was computed with FD threshold for displacement being 0.5. In addition to the frame corresponding to the displaced time point, 1 preceding and 2 succeeding time points were also deleted to reduce the 'spill-over' effect of head movements. Subjects with >10% displaced frames flagged were completely excluded from the analysis as it is likely that such high-level of movement would have had an influence on several volumes. Global signals were not regressed out (3).

Any effects of gender ratio, years of education, head motion, and age between the patient and control groups were regressed out in the analysis. In fact, there were no differences in the gender ratios, though the number of years of education was lower in the patients than controls. We note that the Taiwanese sample included patients with depression in remission while under antidepressant treatment, and thus their scores on the Hamilton Depression Rating Scale (HAMD) assessment were in the low range.

*Xinan***:** All images were acquired on a 3.0-T Siemens Trio MRI scanner using a 16-channel whole-brain coil (Siemens Medical, Erlangen, Germany). High- resolution T1-weighted 3D images were acquired using a magnetization-prepared rapid gradient echo (MPRAGE) sequence (echo time (TE) = 2.52 ms; repetition time (TR) = 1900 ms; inversion time (TI) = 900 ms; flip angle = 9 degrees; slices = 176; thickness = 1.0 mm; resolution matrix = 256×256; voxel size = 1×1×1 mm^3). For each participant, 242 functional images were acquired with a gradient echo type Echo Planar Imaging (EPI) sequence (echo time (TE) = 30 ms; repetition time (TR) = 2000 ms; flip angle = 90 degrees; slices = 32; slice thickness = 3.0 mm; slice gap = 1 mm; resolution matrix = 64×64; voxel size 3.4×3.4×3mm^3). During image acquisition, participants were instructed to keep their eyes closed while keeping their head as still as possible without falling asleep. All participants stayed awake during the MRI imaging as confirmed by the participants after the session.

*Taiwan***:** fMRI scanning was performed at National Yang-Ming University in Taiwan using a 3.0-T Siemens MRI Scanner (Siemens Magnetom Tim Trio, Erlangen, Germany) with a 12-

channel head coil. During the experiments, the participants were instructed to relax with their eyes closed, without falling asleep. After the resting state experiment, participants were asked whether they fell asleep during the resting state scan session, and participants were rescanned if they had fallen asleep during the resting state scan. T2*-weighted images with BOLD contrast were measured using a gradient echo- planar imaging (EPI) sequence (repetition time, TR: 2,500 ms, echo time, TE: 27 ms, field of view, FoV: 220 mm, flip angle: 77 degree, matrix size: 64 x 64, and voxel size: 3.44×3.44×3.40 mm). For each run, 200 EPI volume images were acquired in the anterior and posterior commissure (AC − PC) plane. High-resolution structural T1 images were acquired with three-dimensional (3D) magnetization-prepared rapid gradient-echo sequence (3D-MPRAGE; TR: 2,530 ms, TE: 3.5 ms, TI: 1,100 ms, FoV: 256 mm, and flip angle: 7 degree, 192 sagittal slices, voxel size = 1.0 mm x 1.0 mm 1.0 mm, no gap). For each participant, the whole fMRI scanning lasted about 16 min (T1: 8min, Resting: 8min).

Effective connectivity measurement

Introduction

A classical approach to measuring effective connectivity is dynamic causal modelling (DCM) (7-9). DCM is often used with circuits consisting of a-priori selected brain regions to test hypotheses on the interactions between the considered regions. Here we instead use a network model with simpler assumptions than those typically used in DCM to perform a whole-brain connectivity analysis (10). This allows for the very efficient calculation of maximum-likelihood EC estimates for a large number (94) of nodes, individually for a large cohort of >500 subjects. In this way we target significant EC differences for all existing connections (as determined by DTI) that characterize MDD with FDR correction and without preliminary knowledge, expecting a distributed pattern of abnormal EC links across the brain. Our estimation procedure (10) iteratively optimizes a network model such that it reproduces the empirical cross-covariances between ROIs, which are canonically related to the cross spectral density used in recent studies that apply DCM to resting state fMRI data (11, 12). The proposed model uses an exponential approximation of BOLD autocovariance (locally over a few TRs) and discards very slow-frequency fluctuations. Moreover, by focusing on BOLD fluctuations in the frequency range (0.1-1 Hz), we were able to dispense with a model of haemodynamic

mapping neuronal activity to fMRI signals, as the corresponding time constants are faster (13). Finally, we place positivity constraints on extrinsic or between node connections - in line with known neuroanatomy and previous modeling studies (14). A last simplification compared to DCM includes a fixed (but plausible) form of endogenous neuronal fluctuations (Σ in our model) that were characterized by a single (variance) parameter in each region or node. In spite of these differences, we still borrow the term "effective connectivity" from the DCM literature as our connectivity estimates relate to directional interactions between ROIs in the brain network. This model-based approach has been successfully applied to identify changes in the cortical coordination between rest and movie viewing (15).

Compared to DCM the new method used here (10) is more powerful because it limits the degrees of freedom for each brain region by utilizing a simpler model of each brain region, and because it uses some structural connectivity information from for example diffusion tensor imaging. Further, the new effective connectivity method focuses on transitions between fMRI "activity states" across successive time points (16) and does not include details about hemodynamics like the Balloon model (17). The estimated effective connectivity measures the strengths of causal interactions from one brain area to another, via the proxy of BOLD fluctuations: it provides a single number that lumps together the effects of the strength of the synapse, and neurotransmitter release, etc. The synaptic conductivity interpretation also relates to our earlier neuron-level models in which the synaptic conductivity between modules is a key parameter that specifies how much one module influences another module (18). The new method has the additional advantage that each brain region or module has its own Σ parameter which specifies the variance of the module's activity, which may be related to the intrinsic excitability of the region. In relation to our integrate-and-fire models, the parameter $w+$ that defines the strength of the recurrent collateral synapses within the attractor network (18) may relate to the Σ parameter in the current effective connectivity approach (10), because the local feedback influenced by w+ influences the fluctuations of the activity, for example how readily an area will transition to a high firing rate state.

Within a cortical hierarchy of connectivity (for example from primary visual cortex V1 to the inferior temporal cortex (19)), the forward connections between any pair of cortical areas up through the hierarchy are thought to be stronger than the backprojections based on a wealth

of evidence (20), and there are useful asymmetries in the terminations of the forward and backward projections that facilitate this (20, 21). This ensures that sensory input dominates the processing, rather than imagination. In the present investigation, we follow this lead and refer to the direction in which the effective connectivity is the stronger as the forward direction. The concept of forward relates to function as much as to anatomy. Further, the findings in this paper strengthen the evidence for the use of the term "forward connections", in that the measured ECs described here are what would be expected from our understanding of processing in cortical hierarchies (20).

Overview

The approach used to calculate effective connectivity (EC) follows that described by Gilson et al (10). Effective connectivity measures the individual efficacy of each existing connection between two brain regions, that is, how much one brain region influences another. Our approach provides a signature for each subject in the high-dimensional space of EC connections (>3000), which is then used to investigate differences between MDD patients and healthy controls. The estimated effective connectivity values reflect the combined effects of synaptic efficacies between the regions, the types and concentrations of neurotransmitters in the target regions, etc.

The dynamics for each brain region are described by a multivariate Ornstein-Uhlenbeck process: each region receives fluctuating inputs (white noise) that propagates via the effective connectivity to other nodes, which shapes the correlation pattern at the global level, that is, the functional connectivity (FC). Here the focus is on transitions of fMRI measurements across successive TRs, which have been shown to convey information about conditions such as waking versus sleep (16). The effective connectivity model captures this information via the covariances with non-zero time shifts (spatiotemporal FC) and the resulting EC contains information about directed connectivity. Both EC and the local input variance are optimized such that the model best reproduces statistics of observed fMRI signals measured by the empirical spatiotemporal FC, which are canonically related to the cross spectra used to tune a resting-state DCM (12).

Details about the optimization are provided by Gilson et al (10) for resting-state fMRI data and are summarized next. The skeleton for the EC is provided by structural data obtained using diffusion tensor imaging (DTI), from which we infer the existence of connections. This usefully reduces the number of parameters to estimate and enhances the estimation procedure at the level of individual subjects: from all possible $94^2 = 8836$ connections, we specify that many are not present anatomically as direct projections, so in the model need to optimize only 39% of the possible connections. The DTI connectivity matrix was set to just 0 (no connection) or 1 (for a connection) between the automated anatomical atlas AAL2 94 regions (22), based on the DTI atlas used by Gilson et al (2016). Because DTI may miss inter-hemispheric connections between homotopic regions on the two hemispheres (23, 24), we set these as being present, and allowed the algorithm to tune the strengths of these just as for the other effective connectivities. The automated anatomical atlas version 2 (AAL2) (22) was used to parcellate the brain into 94 regions, because this number of regions provides a suitable number of functional connectivity links without too many degrees of freedom; because its parcellation of the orbitofrontal cortex region which is of special interest in relation to depression (25) has been remade to relate to useful divisions and descriptions; and because it has been found to be useful in related investigations (3). Limiting the number of parameters to estimate in the whole-brain dynamic model is crucial to obtain robust individualized EC estimates. On the other hand, AAL2 corresponds to about 3000 EC parameters (for 39% density), which is a sufficiently rich space to extract complex patterns to differentiate between patients and controls. Our choice aimed to solve this trade-off.

Empirical covariances:

For the resting-state session of each individual, we denote the centered (zero-mean) time series by s_i^t for region $1 \le i \le N$ with time $1 \le t \le T$; the duration is T=180. The zero-lag and 1-lag covariances are calculated as follows:

$$\hat{Q}_{\mathrm{ij}}^0 = \frac{1}{T-2}\sum_{1\le t\le T-1} s_i^t s_j^t \text{ and } \hat{Q}_{\mathrm{ij}}^1 = \frac{1}{T-2}\sum_{1\le t\le T-1} s_i^t s_j^{t+1}. \quad (1)$$

For each subject, we evaluate the time constant τ_x associated with the exponential decay of the autocovariance function $\hat{Q}_{\mathrm{ii}}^{\tau}$ averaged over all regions:

$$\tau_x = \frac{N}{\sum_i \log(\hat{Q}_{ii}^0) - \log(\hat{Q}_{ii}^1)}. \quad (2)$$

Dynamic cortical model:

The model comprised $N = 94$ interconnected cortical regions. The activity x_i of each region is governed by an Ornstein-Uhlenbeck process and evolves depending on the activity of other populations: $\frac{dx_i}{dt} = \frac{-x_i}{\tau_x} + \sum_{j \neq i} C_{ij} x_j + dB_i$. Here, the time constant τ_x corresponds to an exponential decay and is calibrated from the empirical data (see Eq. 2); dB_i is white Gaussian noise with covariance matrix Σ, where the input variances sit on the diagonal and are zero elsewhere. These input fluctuations propagate via the effective connectivity embodied by the matrix C (its skeleton is determined by DTI). All variables x_i have zero mean and the theoretical spatiotemporal covariances are defined by $Q_{ij}^\tau = \langle x_i^t x_j^{t+\tau} \rangle$, where the angular brackets denote averaging over randomness of the inputs; we use two time shifts: $\tau = 0$ and $\tau = 1$ TR.

The mathematical mappings between matrices C, Q^0 and Q^1 are given by Lyapunov equation $JQ^0 + Q^0 J^T + \Sigma = 0$ and $Q^1 = Q^0 \text{expm}(J^T)$, where the Jacobian of the dynamical system $J_{ij} = -\frac{\delta_{ij}}{\tau_x} + C_{ij}$ depends on the mean activity of the network (δ_{ij} is the Kronecker delta); the superscript T denotes the matrix transpose; expm denotes the matrix exponential. These two consistency equations allow for the quick estimation of the predicted FC matrices, without simulating the network.

Parameter estimation procedure:

We tune the model such that its covariance matrices Q^0 and Q^1 reproduce the empirical $\hat{Q}^0$ and $\hat{Q}^1$. We summarize the essential steps of the procedure described in Gilson et al. (2016) that iteratively optimizes the network parameters C and Σ. At each step, the Jacobian J is calculated from the current value of C. Then, the model FC matrices Q^0 and Q^τ are calculated from the consistency equations, using the Bartels-Stewart algorithm to solve the Lyapunov equation. The desired Jacobian update is the matrix $\delta J^T = (Q^0)^{-1}[\delta Q^0 + \delta Q^1 \text{expm}(-J^T)]$, which aims to reduce the FC error between the empirical and model FC, as

determined by the two difference matrices $\delta Q^0 = \hat{Q}^0 - Q^0$ and $\delta Q^1 = \hat{Q}^1 - Q^1$. Finally, the connectivity update is $\delta C_{\mathrm{ij}} = \eta_C \delta J_{\mathrm{ij}}$ for existing connections. We impose non-negativity of the EC values during the optimization. The input variances are tuned according to $\delta \Sigma_{\mathrm{ii}} = -\eta_\Sigma (J \delta Q^0_{\mathrm{ii}} + \delta Q^0_{\mathrm{ii}} J^T)$. We use $\eta_C = 0.0001$ and $\eta_\Sigma = 0.1$.

Normalisation of model estimates

Normalisation of the effective connectivity within each individual was performed by performing the z-score over the matrix elements for each EC matrix within each participant: (EC_{ij} - mean(EC_{ij})) / std(EC_{ij}) for all EC links (performed over existing links corresponding to 1's in the structural connectivity matrix). The aim of this was to enable each participant's data to contribute similarly to the statistics calculated across participants. We note that small effective connectivities will appear in the tables in this paper as negative, but this is only due to the removal of the mean value. All effective connectivity links computed by the algorithm are in fact positive. Any difference between patients and controls that is describe as negative in fact refers only to a decrease of effective connectivity. A similar normalization within each individual was used for the Σ values. An increase of a Σ value can be interpreted as an increase of the variance in an AAL2 region. These normalisations were used for the statistical calculations. Table 1 shows the mean of the EC values not normalized (because a negative EC would have no meaning). Table 2 shows the mean of the Σ values normalized within each participant because this better reflects the statistical values.

Statistical analysis on effective connectivity

Two-tailed, two-sample t-tests were performed on the normalized effective connectivity to identify significantly altered effective connectivity links in patients compared to controls within each imaging centre that provided resting-state fMRI data. The effects of age, gender ratios, head motion and education were regressed within each dataset in this step by general linear models (26, 27). After obtaining the t-test results for each centre, the Liptak-Stouffer z score method (28), which has been described in detail in our previous studies (3, 29, 30), was then used to combine the results from the individual datasets. Further, given that a reasonable

effect size for effective connectivity is in the range .05 to 1 Hz for EC strengths not corrected by the time constant (31), we consider here only EC differences greater than a threshold value of 0.01 (calculated by taking 0.05 and dividing it by the approximate time constant of approximately 4 s).

Robustness of the effective connectivity analyses

The effective connectivity algorithm (10) used here produced robust results, in that with this size of dataset, many links were found with different effective connectivity values in depressed patients vs controls that were significant fully corrected (FDR) across all AAL2 areas. The results were also robust in that when the data set was split into two half splits, the two splits had a correlation for the values of the 3,490 effective connectivity links of $r = 0.26$ with each other ($p<0.0001$) (Fig. S3). Of the 50 effective connectivity links that were significantly different between depressed patients and controls in the full dataset, 6 of the same links were significant in one half split, and 13 in the other half split. The implication is that the results are reasonably reliable when a comparison involving differences between 336 patients and 350 controls.

Validation of the effective connectivity analyses

The effective connectivity approach used here was validated in the sense that many of the key effective connectivity links in the human brain were stronger in the forward direction than the backward direction where this might be expected on the basis of anatomical findings in non-human primates (20, 32, 33). For example, the orbitofrontal cortex receives important forward visual inputs from the temporal lobe visual cortical areas (20, 34), and the effective connectivity was greater in this direction. The forward connectivity is indicated anatomically by connections from superficial pyramidal cells that terminate primarily in layers and 2-3; and backprojections are from deep pyramidal cells that terminate primarily in layer 1 on the apical dendrites of the pyramidal cells, far from the cell bodies, where they have relatively weak effects (20, 21). (Indeed, for the operation of coupled excitatory and inhibitory neuronal pools computational simulations show that modulatory effects that implement attentional biases need to be approximately 0.4 of the strength of the driving connections that convey sensory input

connections (20, 35-37).) The actual values for the ratio of the backprojection to the forward connection effective connectivities shown in Table 1, consistent with this theory, have a mean value of 0.467 for the orbitofrontal cortex backprojections to the temporal lobe cortical areas. (The ratio of these forward to their backprojections in Table 1 is 2.14.) The fact that this baseline value for the effective connectivity in the forward vs backward directions is evident in resting-state fMRI is of considerable interest.

Table S2. The anatomical regions defined in each hemisphere and their label in the automated anatomical labelling atlas AAL2 (22). Column 4 provides a set of possible abbreviations for the anatomical descriptions.

NO.	ANATOMICAL DESCRIPTION	LABEL（aal2.nii.gz)	POSSIBLE ABBREVIATION
1, 2	Precentral gyrus	Precentral	PreCG
3, 4	Superior frontal gyrus, dorsolateral	Frontal_Sup	SFG
5, 6	Middle frontal gyrus	Frontal_Mid	MFG
7, 8	Inferior frontal gyrus, opercular part	Frontal_Inf_Oper	IFGoperc
9, 10	Inferior frontal gyrus, triangular part	Frontal_Inf_Tri	IFGtriang
11, 12	IFG pars orbitalis,	Frontal_Inf_Orb	IFGorb
13, 14	Rolandic operculum	Rolandic_Oper	ROL
15, 16	Supplementary motor area	Supp_Motor_Area	SMA
17, 18	Olfactory cortex	Olfactory	OLF
19, 20	Superior frontal gyrus, medial	Frontal_Sup_Med	SFGmedial
21, 22	Superior frontal gyrus, medial orbital	Frontal_Med_Orb	PFCventmed
23, 24	Gyrus rectus	Rectus	REC
25, 26	Medial orbital gyrus	OFCmed	OFCmed
27, 28	Anterior orbital gyrus	OFCant	OFCant
29, 30	Posterior orbital gyrus	OFCpost	OFCpost
31, 32	Lateral orbital gyrus	OFClat	OFClat
33, 34	Insula	Insula	INS
35, 36	Anterior cingulate & paracingulate gyri	Cingulate_Ant	ACC
37, 38	Middle cingulate & paracingulate gyri	Cingulate_Mid	MCC
39, 40	Posterior cingulate gyrus	Cingulate_Post	PCC
41, 42	Hippocampus	Hippocampus	HIP
43, 44	Parahippocampal gyrus	ParaHippocampal	PHG
45, 46	Amygdala	Amygdala	AMYG
47, 48	Calcarine fissure and surrounding cortex	Calcarine	CAL
49, 50	Cuneus	Cuneus	CUN
51, 52	Lingual gyrus	Lingual	LING
53, 54	Superior occipital gyrus	Occipital_Sup	SOG
55, 56	Middle occipital gyrus	Occipital_Mid	MOG
57, 58	Inferior occipital gyrus	Occipital_Inf	IOG
59, 60	Fusiform gyrus	Fusiform	FFG
61, 62	Postcentral gyrus	Postcentral	PoCG
63, 64	Superior parietal gyrus	Parietal_Sup	SPG
65, 66	Inferior parietal gyrus, excluding supramarginal and angular gyri	Parietal_Inf	IPG
67, 68	SupraMarginal gyrus	SupraMarginal	SMG
69, 70	Angular gyrus	Angular	ANG
71, 72	Precuneus	Precuneus	PCUN
73, 74	Paracentral lobule	Paracentral_Lobule	PCL
75, 76	Caudate nucleus	Caudate	CAU
77, 78	Lenticular nucleus, Putamen	Putamen	PUT
79, 80	Lenticular nucleus, Pallidum	Pallidum	PAL
81, 82	Thalamus	Thalamus	THA
83, 84	Heschl’s gyrus	Heschl	HES
85, 86	Superior temporal gyrus	Temporal_Sup	STG
87, 88	Temporal pole: superior temporal gyrus	Temporal_Pole_Sup	TPOsup
89, 90	Middle temporal gyrus	Temporal_Mid	MTG
91, 92	Temporal pole: middle temporal gyrus	Temporal_Pole_Mid	TPOmid
93, 94	Inferior temporal gyrus	Temporal_Inf	ITG

Table S3. Correlations between the effective connectivity links and the depression severity scores, assessed by the duration of the illness. The r value shows the correlation between the effective connectivity and the illness duration calculated just in patients. The z value show the difference in effective connectivity for each link for the comparison depression group – control group, and its p value follows in column 7. Thus for example for the OFCant to temporal pole effective connectivity link, this shows a greater decrease with an increase of illness duration. In addition, correlations of some of the functional connectivities with the Beck Depression Inventory (BDI), the Hamilton Depression rating scale (HAM-D), and the Hamilton Anxiety rating scale (HAM-A) are shown.

Region 1	Region 2	clinical variables	r value	p value of r	z value	p value of z
Precentral_R	Precentral_L	Illness duration	0.108	0.0491	3.320	9.01E-04
OFCant_L	Temporal_Inf_L	Illness duration	-0.135	0.0123	-3.334	8.55E-04
Temporal_Pole_Mid_L	Temporal_Pole_Mid_R	Illness duration	0.121	0.0266	3.300	9.67E-04
Frontal_Mid_2_L	Frontal_Sup_Medial_R	Illness duration	-0.134	0.0144	-3.700	2.15E-04
Olfactory_L	OFCmed_L	HAMA	-0.139	0.0111	-3.410	6.49E-04
Olfactory_R	OFCmed_R	HAMA	-0.112	0.0412	-3.255	1.14E-03
OFCpost_L	OFClat_L	HAMD	0.120	0.0287	3.466	5.29E-04
Temporal_Pole_Mid_L	Temporal_Pole_Mid_R	HAMD	0.111	0.0425	3.300	9.67E-04
OFCmed_R	Temporal_Pole_Mid_R	HAMD	-0.125	0.0226	-4.014	5.98E-05
Amygdala_L	OFCpost_L	BDI	-0.138	0.0353	-3.914	9.09E-05
Cingulate_Post_R	Temporal_Pole_Mid_R	BDI	0.130	0.0473	3.611	3.05E-04
ParaHippocampal_R	Occipital_Mid_R	BDI	-0.158	0.0153	-3.328	8.74E-04

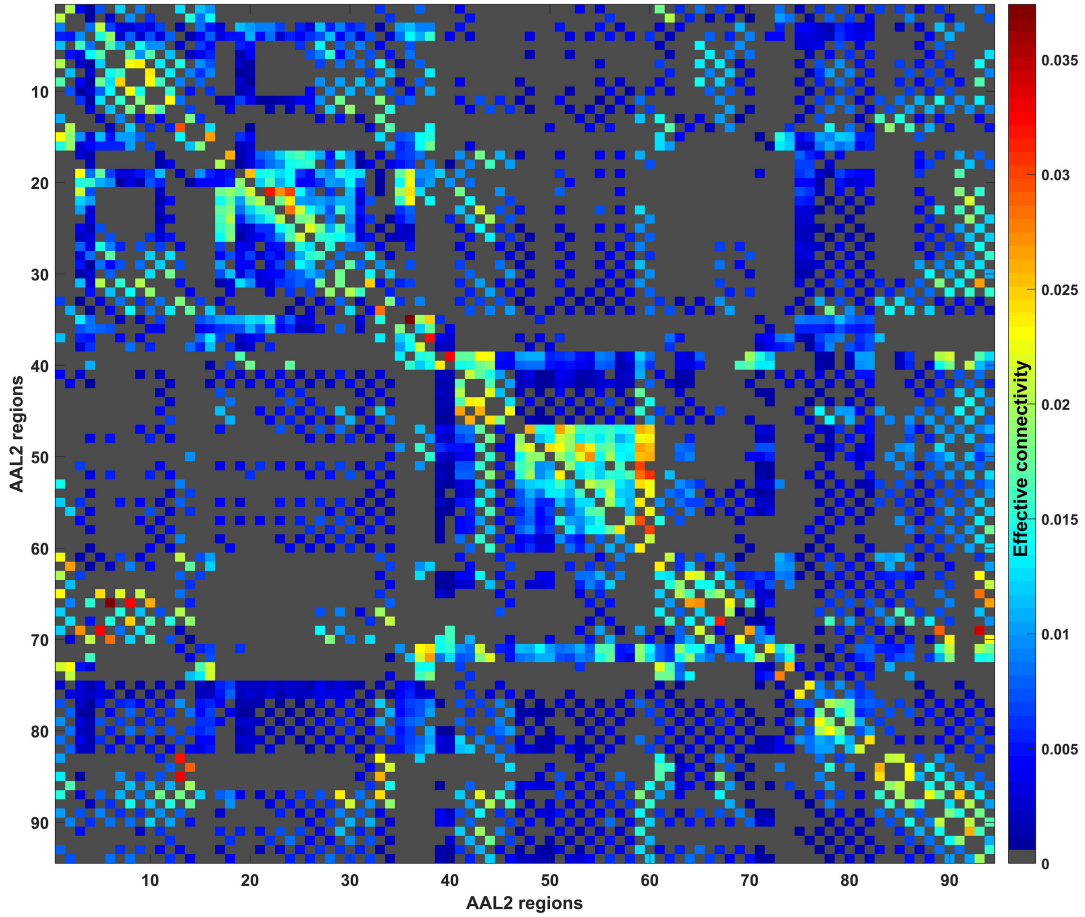

Figure S1. The matrices of effective connectivity of healthy controls. The axes are the AAL2 areas, shown numbered and with their names in Table S1. The effective connectivity matrix has the index j for the columns and the index i for the rows. The matrix is thus non-symmetric, and the effective connectivity is always from j to i. The effective connectivity between any pair of links is shown in one direction in the upper right of the matrix, and in the opposite direction in the lower left. The table shows for example that AAL2 areas 3-12 which include the inferior frontal gyrus reading along the horizontal axis tend to have strong forward effective connectivity to angular and supramarginal gyri areas 67-70 reading on the vertical axis.

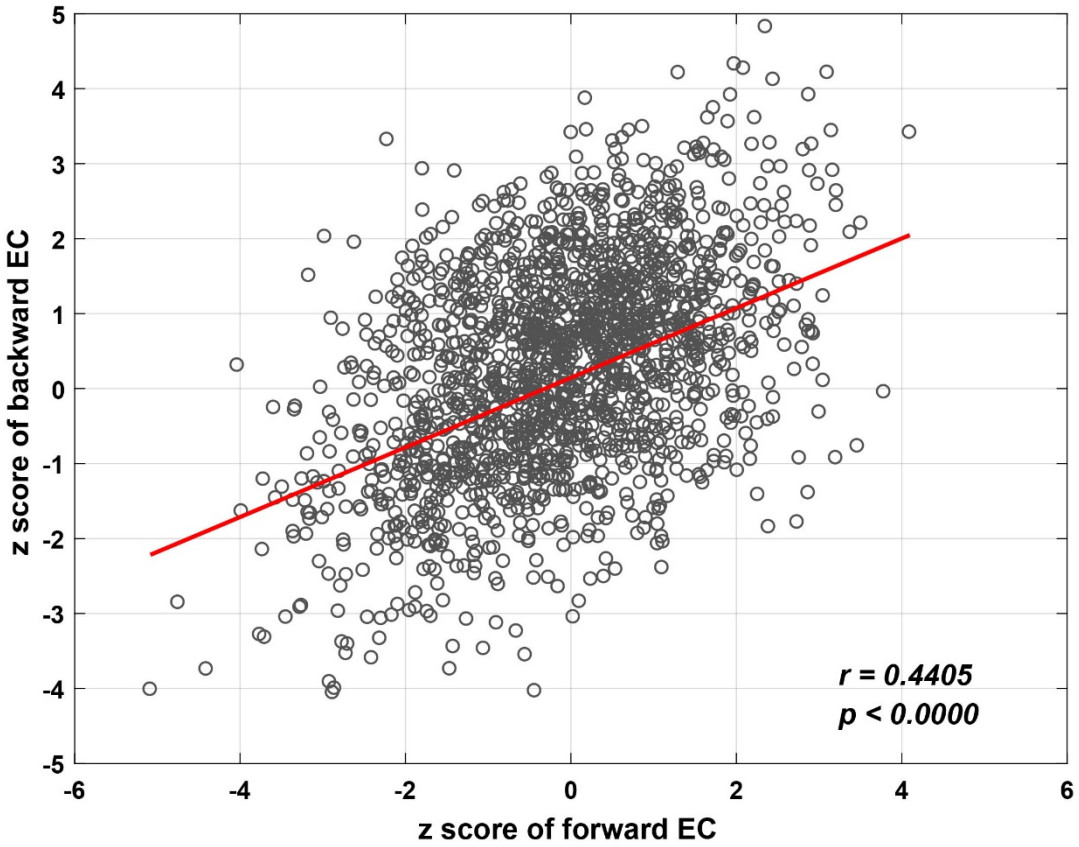

Figure S2. Comparison of the difference between patients with major depressive disorder and controls in the strength of the effective connectivity in the forwards vs the backwards direction. The scattergram shows for every link the z value for the difference between patients and controls in the forward direction on the abscissa and the backward direction on the ordinate. For this diagram, forwards indicates only i to j in the effective connectivity matrix shown in Fig. S1, and backwards indicates j to i. The linear regression line and correlation value show that if a link is weaker in one direction in patients, it is likely to be weaker in the opposite direction; and that if a link is stronger in patients in one direction, it is likely to be stronger in the other direction too.

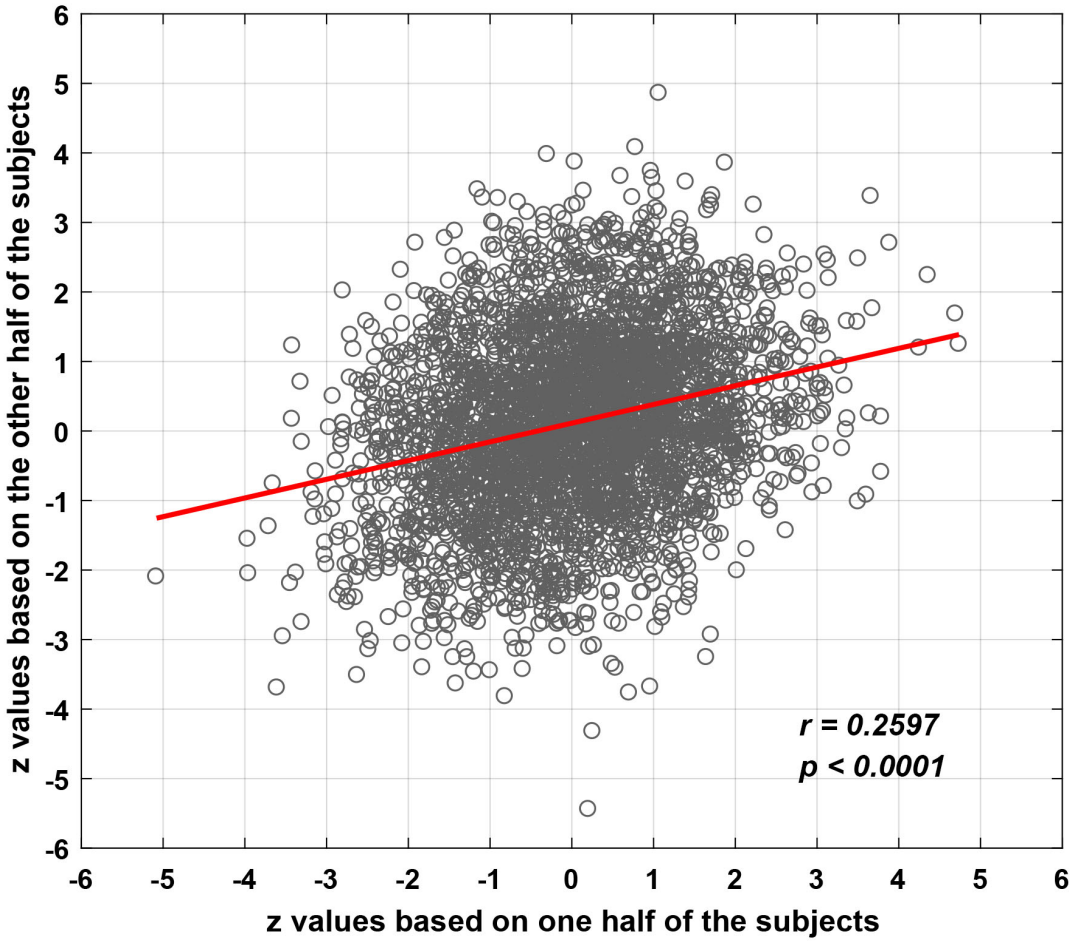

Fig. S3. Comparison of effective connectivity differences between patients and controls when the dataset was split into two halves. The two splits had a correlation for the values of the 3,490 effective connectivity links of r = 0.70 with each other ($p = 1.1 \times 10^{-251}$).

Effects of medication

Although it was not a primary aim of this investigation, and following a suggestion, the effects of medication were assessed by comparing the functional connectivity in 125 patients not receiving medication, and 157 patients receiving medication. The medication consisted in most cases of selective serotonin reuptake inhibitors (SSRIs) including fluoxetine, paroxetine, sertraline, citalopram and escitalopram; or serotonin-norepinephrine reuptake inhibitors (SNRIs) such as venflaxine, or a tetracyclic antidepressant such as mirtazepine. The overall pattern of functional connectivity differences between patients and controls is similar for the unmedicated and the medicated subgroups of patients (Fig. S4), providing evidence that the main differences between patients and controls shown in Fig. 1 were found in depressed

patients whether or not they were receiving medication. This was shown by a correlation of 0.68 ($p<10^{-10}$) between the functional connectivities across the whole brain in medicated and unmedicated patients. Further, the changes shown in Table 1 for the differences between all the patients and controls were similar, with similar signs, in both the medicated and the unmedicated patients vs the controls.

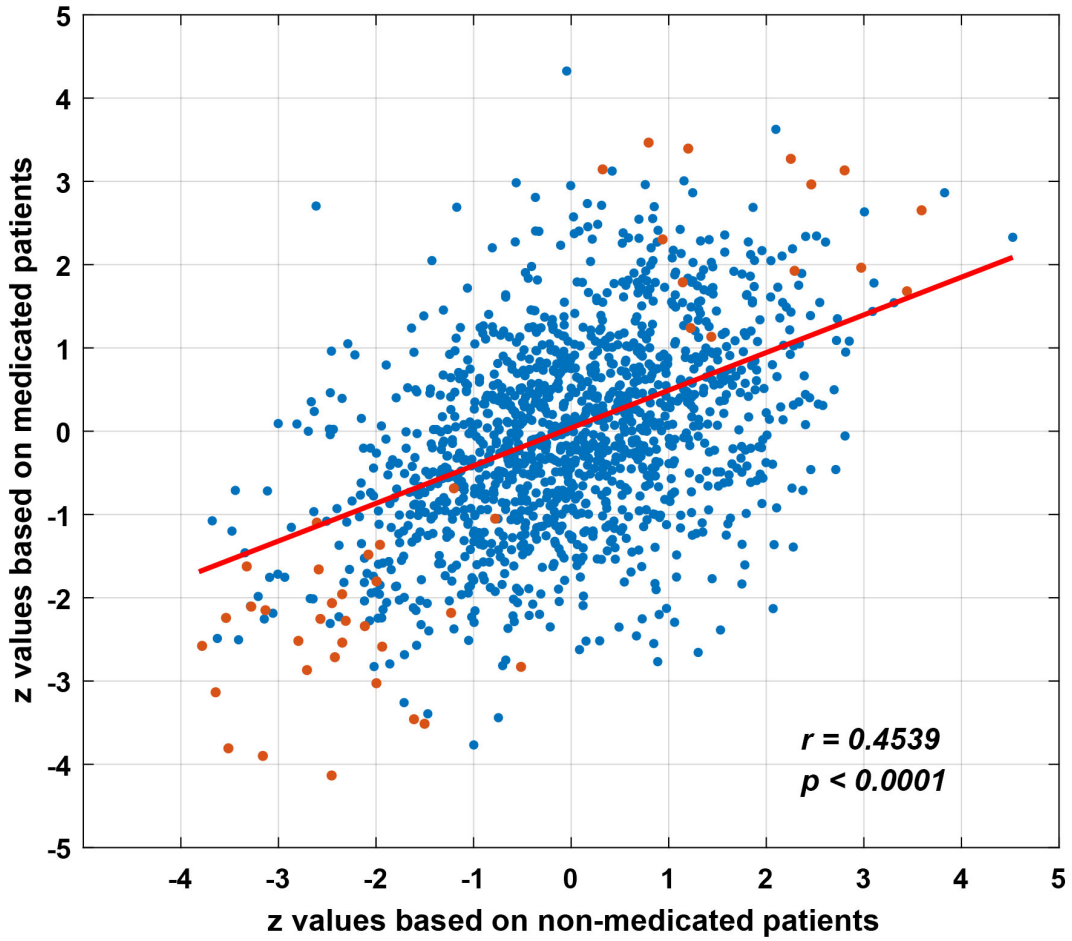

Fig. S4. There is a correlation between the strengths of the functional connectivity links in the non-medicated patients and of the links in the medicated population, showing that a similar pattern of differences from controls is found in medicated and un-medicated patients. The bold points are for links that were significantly different between all patients and controls. The faint points show the links between all other AAL2 areas. The r value is for all links. The red points indicate links included in Table 1.

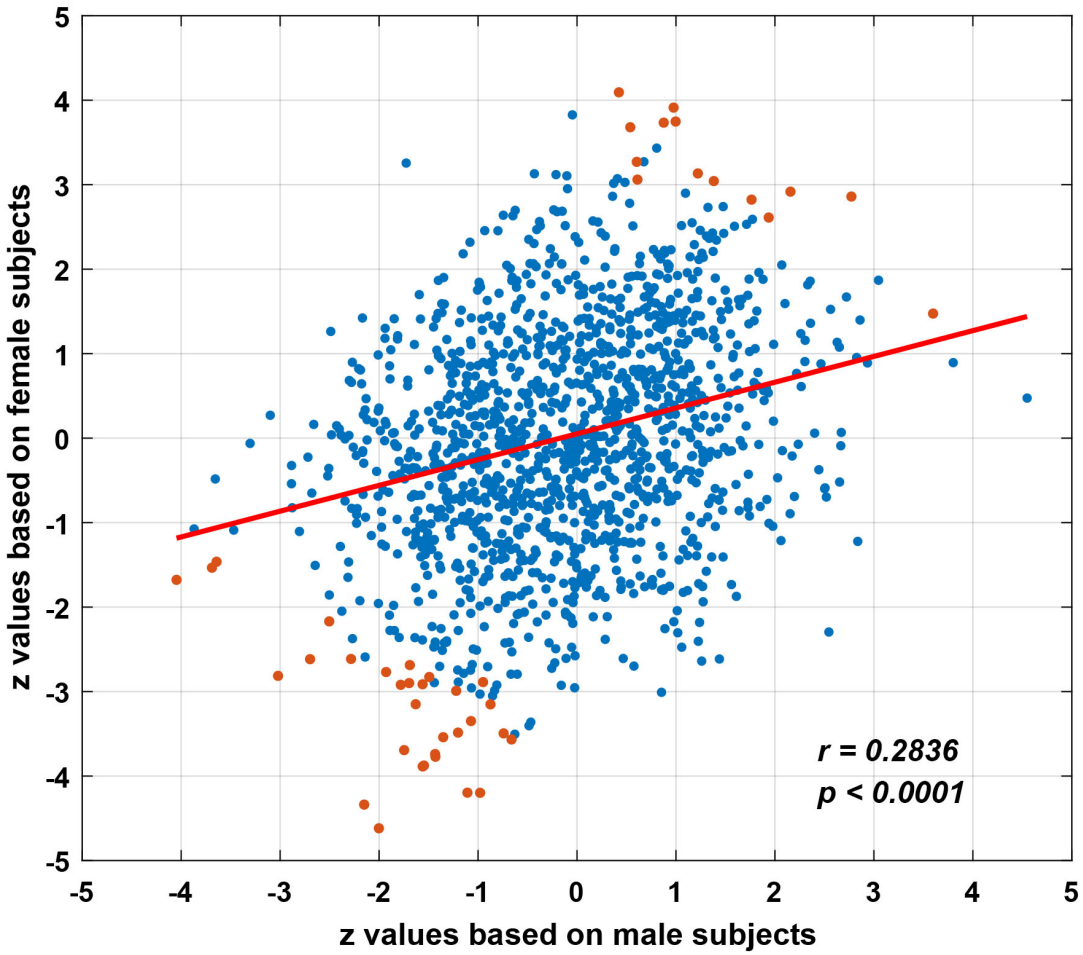

Fig. S5. Comparison of effective connectivity differences between patients and controls when the dataset was split into two groups, male and female. The red points indicate links included in Table 1.

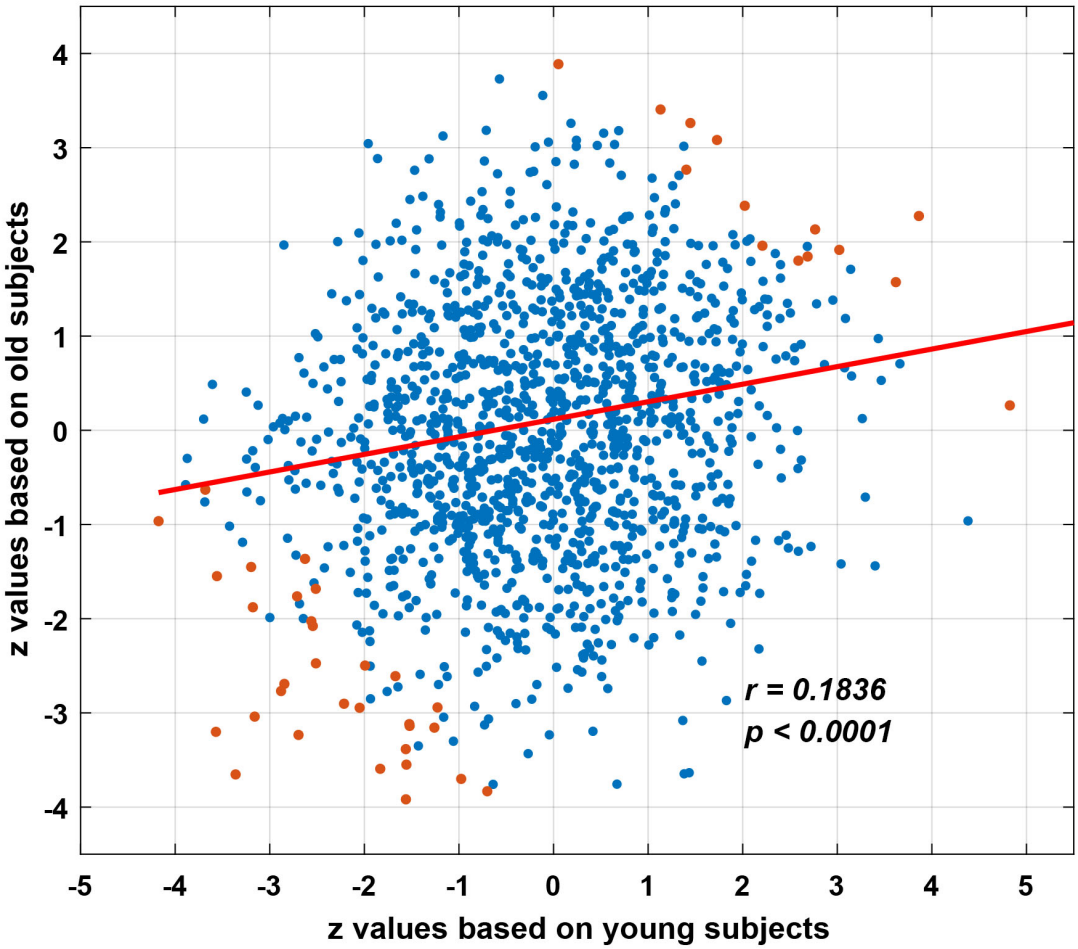

Fig. S6. Comparison of effective connectivity differences between patients and controls when the dataset was split into two groups, younger and older. The red points indicate links included in Table 1.

Supplemental References

1. Hamilton M (1960): A rating scale for depression. *J Neurol Neurosurg Psychiatry*. 23:56-62.
2. Beck AT, Beamesderfer A (1974): Assessment of depression: the depression inventory. *Mod Probl Pharmacopsychiatry*. 7:151-169.
3. Cheng W, Rolls ET, Qiu J, Liu W, Tang Y, Huang CC, et al. (2016): Medial reward and lateral non-reward orbitofrontal cortex circuits change in opposite directions in depression. *Brain*. 139:3296-3309.
4. Chao-Gan Y, Yu-Feng Z (2010): DPARSF: a MATLAB toolbox for "pipeline" data analysis of resting-state fMRI. *Front Syst Neurosci*. 4:13.
5. Friston KJ, Williams S, Howard R, Frackowiak RS, Turner R (1996): Movement-related effects in fMRI time-series. *Magn Reson Med*. 35:346-355.
6. Power JD, Mitra A, Laumann TO, Snyder AZ, Schlaggar BL, Petersen SE (2014): Methods to detect, characterize, and remove motion artifact in resting state fMRI. *Neuroimage*. 84:320-341.
7. Valdes-Sosa PA, Roebroeck A, Daunizeau J, Friston K (2011): Effective connectivity: influence, causality and biophysical modeling. *Neuroimage*. 58:339-361.
8. Bajaj S, Adhikari BM, Friston KJ, Dhamala M (2016): Bridging the Gap: Dynamic Causal Modeling and Granger Causality Analysis of Resting State Functional Magnetic Resonance Imaging. *Brain Connect*.
9. Friston K (2009): Causal modelling and brain connectivity in functional magnetic resonance imaging. *PLoS Biol*. 7:e33.
10. Gilson M, Moreno-Bote R, Ponce-Alvarez A, Ritter P, Deco G (2016): Estimation of directed effective connectivity from fMRI functional connectivity hints at asymmetries in the cortical connectome. *PLoS Comput Biol*. 12:e1004762.
11. Frassle S, Lomakina EI, Razi A, Friston KJ, Buhmann JM, Stephan KE (2017): Regression DCM for fMRI. *Neuroimage*. 155:406-421.
12. Friston KJ, Kahan J, Biswal B, Razi A (2014): A DCM for resting state fMRI. *Neuroimage*. 94:396-407.
13. Friston KJ (2002): Bayesian estimation of dynamical systems: an application to fMRI. *Neuroimage*. 16:513-530.
14. Marreiros AC, Kiebel SJ, Friston KJ (2008): Dynamic causal modelling for fMRI: a two-state model. *Neuroimage*. 39:269-278.
15. Gilson M, Deco G, Friston K, Hagmann P, Mantini D, Betti V, et al. (2017): Effective connectivity inferred from fMRI transition dynamics during movie viewing points to a balanced reconfiguration of cortical interactions. *bioRxiv*.110015.
16. Mitra A, Snyder AZ, Tagliazucchi E, Laufs H, Raichle ME (2015): Propagated infra-slow intrinsic brain activity reorganizes across wake and slow wave sleep. *Elife*. 4:e10781.
17. Friston KJ, Mechelli A, Turner R, Price CJ (2000): Nonlinear responses in fMRI: the Balloon model, Volterra kernels, and other hemodynamics. *Neuroimage*. 12:466-477.
18. Rolls ET, Webb TJ, Deco G (2012): Communication before coherence. *Eur J Neurosci*. 36:2689-2709.
19. Rolls ET (2012): Invariant visual object and face recognition: neural and computational

bases, and a model, VisNet. *Front Comput Neurosci*. 6, 35:1-70.
20. Rolls ET (2016): *Cerebral Cortex: Principles of Operation*. Oxford: Oxford University Press.
21. Pandya DN, Seltzer B, Petrides M, Cipolloni PB (2015): *Cerebral Cortex: Architecture, Connections, and the Dual Origin Concept*. New York: Oxford University Press.
22. Rolls ET, Joliot M, Tzourio-Mazoyer N (2015): Implementation of a new parcellation of the orbitofrontal cortex in the automated anatomical labeling atlas. *Neuroimage*. 122:1-5.
23. Messe A, Rudrauf D, Benali H, Marrelec G (2014): Relating structure and function in the human brain: relative contributions of anatomy, stationary dynamics, and non-stationarities. *PLoS Comput Biol*. 10:e1003530.
24. Hagmann P, Cammoun L, Gigandet X, Meuli R, Honey CJ, Wedeen VJ, et al. (2008): Mapping the structural core of human cerebral cortex. *PLoS Biol*. 6:e159.
25. Rolls ET (2016): A non-reward attractor theory of depression. *Neurosci Biobehav Rev*. 68:47-58.
26. Barnes J, Ridgway GR, Bartlett J, Henley SM, Lehmann M, Hobbs N, et al. (2010): Head size, age and gender adjustment in MRI studies: a necessary nuisance? *Neuroimage*. 53:1244-1255.
27. Di Martino A, Yan CG, Li Q, Denio E, Castellanos FX, Alaerts K, et al. (2014): The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism. *Mol Psychiatry*. 19:659-667.
28. Liptak T (1958): On the combination of independent tests. *Magyar Tud Akad Mat Kutato Int Kozl*. 3:171-197.
29. Cheng W, Rolls ET, Gu H, Zhang J, Feng J (2015): Autism: reduced functional connectivity between cortical areas involved in face expression, theory of mind, and the sense of self. *Brain*. 138:1382-1393.
30. Cheng W, Palaniyappan L, Li M, Kendrick KM, Zhang J, Luo Q, et al. (2015): Voxel-based, brain-wide association study of aberrant functional connectivity in schizophrenia implicates thalamocortical circuitry. *NPJ Schizophr*. 1:15016.
31. Stephan KE, Penny WD, Moran RJ, den Ouden HE, Daunizeau J, Friston KJ (2010): Ten simple rules for dynamic causal modeling. *Neuroimage*. 49:3099-3109.
32. Markov NT, Ercsey-Ravasz M, Van Essen DC, Knoblauch K, Toroczkai Z, Kennedy H (2013): Cortical high-density counterstream architectures. *Science*. 342:1238406.
33. Markov NT, Ercsey-Ravasz MM, Ribeiro Gomes AR, Lamy C, Magrou L, Vezoli J, et al. (2014): A weighted and directed interareal connectivity matrix for macaque cerebral cortex. *Cereb Cortex*. 24:17-36.
34. Rolls ET (2014): *Emotion and Decision-Making Explained*. Oxford: Oxford University Press.
35. Deco G, Rolls ET (2005): Neurodynamics of biased competition and co-operation for attention: a model with spiking neurons. *J Neurophysiol*. 94:295-313.
36. Deco G, Rolls ET (2005): Attention, short-term memory, and action selection: a unifying theory. *Prog Neurobiol*. 76:236-256.
37. Deco G, Rolls ET (2004): A neurodynamical cortical model of visual attention and invariant object recognition. *Vision Res*. 44:621-644.