Prediction of clinical depression scores and detection of changes in whole-brain using resting-state functional MRI data with partial least squares regression

Kosuke Yoshida[1,2]*, **Yu Shimizu[3]**, **Junichiro Yoshimoto[4]**, **Masahiro Takamura[5]**, **Go Okada[5]**, **Yasumasa Okamoto[5]**, **Shigeto Yamawaki[5]**, **Kenji Doya[2]**

1 Graduate School of Informatics, Kyoto University, Kyoto, Japan, **2** Neural Computation Unit, Okinawa Institute of Science and Technology Graduate University, Okinawa, Japan, **3** Integrated Open System Unit, Okinawa Institute of Science and Technology Graduate University, Okinawa, Japan, **4** Graduate School of Information Science, Nara Institute of Science and Technology, Nara, Japan, **5** Department of Psychiatry and Neurosciences, Hiroshima University, Hiroshima, Japan

* kosuke.yoshida@oist.jp

Abstract

In diagnostic applications of statistical machine learning methods to brain imaging data, common problems include data high-dimensionality and co-linearity, which often cause over-fitting and instability. To overcome these problems, we applied partial least squares (PLS) regression to resting-state functional magnetic resonance imaging (rs-fMRI) data, creating a low-dimensional representation that relates symptoms to brain activity and that predicts clinical measures. Our experimental results, based upon data from clinically depressed patients and healthy controls, demonstrated that PLS and its kernel variants provided significantly better prediction of clinical measures than ordinary linear regression. Subsequent classification using predicted clinical scores distinguished depressed patients from healthy controls with 80% accuracy. Moreover, loading vectors for latent variables enabled us to identify brain regions relevant to depression, including the default mode network, *the right superior frontal gyrus*, and *the superior motor area*.

Introduction

Advances in analyzing large datasets with machine learning algorithms promote their application in medical diagnosis. In particular, their use in objective diagnosis of psychiatric disorders using brain imaging and other biological data is now being actively studied [1]. A major challenge in applying statistical machine learning algorithms to brain imaging or genetic data is the high dimensionality of the input variables, such as the number of voxels and the number of possible genetic polymorphisms. Even though algorithms such as support vector machine (SVM) and L1-regularized classifiers (LASSO) manage the issue of high-dimensionality, the problem of co-linearity in brain imaging data remains. Neural activities in nearby voxels or in the same functional network are highly correlated, which makes the results of commonly used

regression or classification tools unreliable. In this paper, we propose the use of partial least squares (PLS) regression [2–7] with multiple clinical measures to address this problem. We use resting-state functional magnetic resonance imaging (rs-fMRI) data and clinical measures from clinically depressed patients and healthy control subjects to obtain low-dimensional representations of symptoms and brain activities, and we use them to predict depression-related clinical measures and thereafter, to classify subjects.

Use of rs-fMRI is gaining attention in diagnosis of psychiatric disorders because it makes few cognitive demands in measurements and because it can be applied to multiple disorders [8]. In depressed patients, functional connectivities (FCs) between brain areas estimated using rs-fMRI show distributed changes throughout the entire brain [9–12]. Zeng et al. (2012) [13] demonstrated that $\sim 94\%$ of 53 subjects could be correctly classified as patients or healthy controls using FCs and linear SVM, and they reported that the majority of discriminating FCs were distributed within or across the default mode network, the affective network, visual cortical areas, and the cerebellum. While the aforementioned study sought to discern differences between patients and healthy controls in a binary manner, Zhang et al. (2011) [14] tried to predict clinical measures of the Beck Depression Inventory II (BDI-II) [15] by regressing fMRI signals acquired during a face-watching task. They showed that true and predicted BDI-II were significantly correlated ($r = 0.55$) and using the standard threshold of 14 for the predicted BDI-II, 89% of the automated diagnoses agreed with those of psychiatrists.

Clinical depression is characterized by multiple, related symptoms [16]. There are various clinical measures for assessing symptoms, such as the Snaith-Hamilton Pleasure Scale (SHAPS) [17] for anhedonia and Positive and Negative Affect Schedule (PANAS) [18] for altered mood. In addition, the age of subjects is important for diagnosis since aging increases the risk of depression in general [19].

Here, we consider a two-step approach which predicts multiple measures of clinical depression from rs-fMRI in the first step, and then uses results of the first step for diagnosis. For the first step, we train a regression model to predict BDI-II, SHAPS, PANAS(n), and age from functional connectivity data. Although this could be done using ordinary least squares regression, in order to tackle the issue of high-dimensionality and co-linearity of the input, we explore the use of partial least squares (PLS) regression [2–7], which maps input and output variables to low-dimensional spaces so that the covariance of data in the latent spaces is maximized. We compare the classification performance of the two-step approach using PLS regression with other classification methods. Thereafter, we consider the use of subject age by testing (i) a model with age as a response variable (output-age model), (ii) a model with age as a predictor (input-age model), and (iii) a model that does not consider age (no-age model).

This paper further develops the basic idea presented in [20] to overcome limitations of linear methods and perform objective diagnosis. In section 2, we illustrate the details of rs-fMRI and clinical measures for subjects. Section 3 provides the mathematical basis of PLS and its kernel variants. In addition, it is extended to classification models for the purpose of objective diagnosis. In section 4, we illustrate the efficacy of our application in predicting clinical measures, discriminating between patients and healthy controls, and interpreting derived coefficients. Finally, we offer our conclusions and discuss future work in section 5.

Data set

This study was approved by the Human Subjects Research Review Committee at the Okinawa Institute of Science of Technology, as well as the Research Ethics Committee of Hiroshima University (permission nr. 172). Written and informed consent was obtained from all subjects participating in the study.

Subjects

58 patients (age 26–73, average 42.8 ± 11.9, 33 female) with major depression disorder were recruited by the Psychiatry Department of Hiroshima University and collaborating medical institutions, based on the Mini-international neuropsychiatric interview (M.I.N.I [21]), which enables doctors to identify psychiatric disorders, according to the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV [22]). As a healthy control group, 65 subjects (ages 20–66, average 34.8 ± 13.0, 28 female) with no history of mental or neurological disease, were recruited via advertisements in local newspapers.

Clinical measures

The following interview- and questionnaire-based measures are used for determination of disease presence and quantification of the severity of two primary symptoms we wish to predict, namely, anhedonia (loss of motivation, loss of pleasure, etc.) and negative mood (low mood, guilty feelings, suicidal thoughts, etc.).

Beck Depression Inventory II (BDI-II).　This measure evaluates the presence and severity of depression based on a self-report questionnaire [15]. Subjects are asked to answer 21 questions about feelings of punishment or guilt, suicidal thoughts, etc. Each answer is scored with a value between 0 and 3, with 3 being the most serious. High scores indicate severe symptoms. The standardized score of ≥ 14 indicates that a subject is suffering from depression.

Snaith-Hamilton Pleasure Scale (SHAPS).　This measure was developed to evaluate the level of anhedonia [17]. Subjects are asked to answer 14 questions about hedonic capacity, with scores between 1 and 4. High scores indicate more severe anhedonia.

Positive and Negative Affect Schedule (PANAS).　This widely used measure evaluates positive and negative moods of subjects [18, 23]. In this study, we considered only scores related to negative mood items. This measure is generally known as PANAS(n). Subjects are asked to respond to 10 questions about their moods, with answers between 0 and 5. The sum of all scores indicates the strength of their negative moods. Due to an evaluation issue, one subject's response could not be assessed, so that score was replaced with the mean of the remaining subjects.

Table 1 summarizes scores exhibited for each measure by each group in our study [20]. Although most patients showed both anhedonia and negative mood, some exhibited only one trait. Correspondingly, the scores of the BDI-II, SHAPS, and PANAS(n) are highly, but not completely correlated. As decreased mental function results from aging, the age of the subjects is expected to correlate with BDI-II, SHAPS, and PANAS(n) as well.

We verified these correlations by calculating the correlation coefficients (Table 2) [20]. Strong correlations between clinical measures are reflected in coefficients above 0.7. Weaker correlations between age and individual clinical measures were around 0.3. In our regression model, BDI-II, SHAPS, PANAS(n), and age of each subject are considered as responses in

Table 1. Mean (± standard deviation) of clinical measures [20].

	Controls	Patients
Number of subjects	65	58
Age	34.8 (±13.0)	42.8 (±11.9)
BDI-II	6.92 (±5.9)	30.9 (±9.0)
SHAPS	23.3 (±6.2)	37.8 (±5.5)
PANAS(n)	8.5 (±6.4)	25.1 (±7.9)

Table 2. The Pearson's correlation coefficients between the clinical measures and the subjects' age [20].

	Age	BDI-II	SHAPS	PANAS(n)
BDI-II	0.2451	-	0.7883	0.8005
SHAPS	0.3221	0.7883	-	0.7497
PANAS(n)	0.2480	0.8005	0.7497	-

order to correct for their natural correlation, resulting from functional connectivity. We will show that the introduction of subject age as an output rather than as an input is beneficial with respect to classification accuracy.

Functional connectivity of resting-state fMRI

Functional MRI measurements were acquired on a 3T GE Signa HDx scanner with a 2D EP/GR (TR = 3s, TE = 27ms, FA = 90deg, matrix size 64x64x32, voxel size 4x4x4 mm, no gap, interleaved). Subjects were instructed to lie with their eyes open, to think of nothing in particular, and to remain awake. They are also instructed to refrain from taking caffeine, nicotine, and alcohol in the day of experiment.

For each subject, acquired images were processed with SPM8 (Wellcome Trust Centre for Neuroimaging, UCL, London) following standard procedures. We first perform slice timing correction, motion correction, co-registration to anatomical MRI, normalization with standard brain and smoothing (Gaussian of full-width at half-maximum 8mm). We confirmed that there were no significant differences in six motion parameters between two diagnostic groups in order to reject a possible effect of spurious functional connectivity due to head motion [24, 25]. Voxels were assigned to 116 brain regions, according to the automatic anatomical labeling atlas (AAL) [26]. Mean activation time series in each brain region were obtained by averaging MRI signal time series over all voxels assigned to each region. Finally, functional connectivity between each pair of regions was computed as the cross correlation of the corresponding time-series.

Methods

Partial least squares (PLS) regression is a method for modeling a relationship between two sets of multivariate data via a latent space, and of performing least squares regression in that space. PLS can handle high-dimensional co-linear datasets because of its underlying assumption that the two datasets are generated by a small number of latent components. In this process, latent components are formed by maximizing the covariance between the two datasets.

Partial Least Squares Regression (PLS)

PLS models a linear relationship between two blocks of variables $\{\mathbf{x}_i\}_{i=1}^{n} \in \mathbb{R}^p$ and $\{\mathbf{y}_i\}_{i=1}^{n} \in \mathbb{R}^q$. In the following parts, $X = (\mathbf{x}_1, \ldots, \mathbf{x}_n)^T$ represents the $(n \times p)$ predictor matrix and $Y = (\mathbf{y}_1, \ldots, \mathbf{y}_n)^T$ represents the $(n \times q)$ response matrix. This procedure obtains L latent components as $\{\mathbf{t}_i\}_{i=1}^{L}$ and $\{\mathbf{u}_i\}_{i=1}^{L}$. This technique assumes following decomposition:

$$X = TP^T + F_x \tag{1}$$

$$Y = UQ^T + F_y, \tag{2}$$

where both $T = (\mathbf{t}_1, \ldots, \mathbf{t}_L)$ and $U = (\mathbf{u}_1, \ldots, \mathbf{u}_L)$ are the $(n \times L)$ matrices of L latent components corresponding to X and Y, respectively. The $(p \times L)$ matrix P and the $(q \times L)$ matrix Q are

loadings and the $(n \times p)$ matrix F_x and the $(n \times q)$ matrix F_y are the matrices of residuals. Since our objective is to perform least squares regression in a low-dimensional latent space, the underlying assumption is that the latent component $\mathbf{u}_i$ can be well predicted from $\mathbf{t}_i$ from a relation such as:

$$U = TD, \tag{3}$$

where D is the $(L \times L)$ matrix. We need to maximize the covariance between $\mathbf{t}_i$ and $\mathbf{u}_i$ to satisfy the above assumption.

Our objective criterion is

$$\max_{\mathbf{t},\mathbf{u}} \mathsf{COV}(\mathbf{t}, \mathbf{u}) = \max_{\mathbf{w},\mathbf{c}} \mathsf{COV}(X\mathbf{w}, Y\mathbf{c}), \tag{4}$$

where $\mathbf{w} \in \mathbb{R}^p$ and $\mathbf{c} \in \mathbb{R}^q$ are weight vectors for projection into the latent components.

After extracting the latent component, the observation matrices X and Y are deflated by subtracting their rank-one approximation. It is important to stress the asymmetry scheme, i.e. that Y is deflated based on $\mathbf{t}$, in the case of regression. By repeating the above procedures L times, we obtain the weight matrices $W = (\mathbf{w}_1, \ldots, \mathbf{w}_L)$ and $C = (\mathbf{c}_1, \ldots, \mathbf{c}_L)$.

Finally, the relation in the original data space is expressed by

$$Y = XB + E, \tag{5}$$

where B is the $(p \times q)$ matrix of regression coefficients and E is the $(n \times q)$ matrix of residuals.

Plugging the relationship $B = W(P^T W)^{-1} C^T$ [27, 28] into Eq (5), we obtain a different representation of Y as:

$$\hat{Y} = XB \tag{6}$$

$$= XW(P^T W)^{-1} C^T \tag{7}$$

$$= XX^T U(T^T XX^T U)^{-1} T^T Y. \tag{8}$$

The final transformation is derived from the following equalities [29],

$$W = X^T U, \tag{9}$$

$$P = X^T T, \tag{10}$$

$$C = Y^T T. \tag{11}$$

Note that $\mathbf{t}_i^T \mathbf{t}_j = \delta_{ij}$ (the Kronecher delta) takes the values 1 for $i = j$ and 0 for $i \neq j$ as a consequence of the algorithm.

In general, B is obtained from a centered training dataset. The response $\mathbf{y}_{new}$ for a new subject $\mathbf{x}_{new}$, referred to as test dataset, is then estimated as follows:

$$\mathbf{y}_{new} = \bar{\mathbf{y}} + B^T(\mathbf{x}_{new} - \bar{\mathbf{x}}), \tag{12}$$

where $\bar{\mathbf{y}}$ and $\bar{\mathbf{x}}$ represent the mean predictor and response in the training dataset, respectively. A schematic outline of PLS is illustrated in Fig 1 and S1 Appendix.

Kernel Partial Least Squares Regression (KPLS)

Linear PLS is easily extended to nonlinear regression using a kernel trick [28, 30].
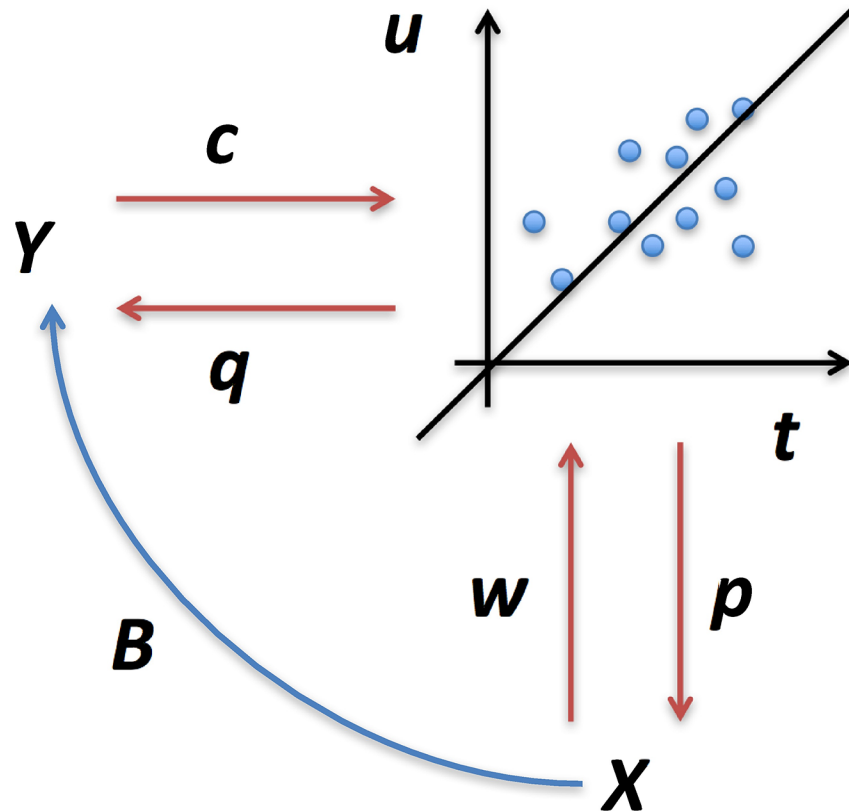
Fig 1. Schematic illustration of partial least squares regression. Two blocks of data, *X* and *Y*, are projected by **w** and **c** onto latent components, **t** and **u**, and least squares regression is performed. **p** and **q** represent loading vectors.

Let $\phi : \mathbb{R}^p \rightarrow \mathcal{H}$ be a nonlinear transformation of the predictor, $\mathbf{x} \in \mathbb{R}^p$, into a feature vector, $\phi(\mathbf{x}) \in \mathcal{H}$, where $\mathcal{H}$ is a high-dimensional feature space. Define a Gram matrix K as inner products of points in feature space, i.e., $K = \Phi\Phi^T$, where $\Phi = (\phi(\mathbf{x}_1), \ldots, \phi(\mathbf{x}_n))^T$ represents the predictor matrix in feature space. In general, the number of columns of Φ is so large that with the explicit form of Φ, we can not perform the same procedure as in the linear case. However, due to the kernel trick, the explicit form of Φ becomes unnecessary.

The deflation procedure is performed as follows:

$$K \quad \leftarrow \quad (I_n - \mathbf{t}\mathbf{t}^T)K(I_n - \mathbf{t}\mathbf{t}^T) \tag{13}$$

$$Y \quad \leftarrow \quad Y - \mathbf{t}\mathbf{t}^T Y, \tag{14}$$

where I_n represents an n-dimensional identity matrix.

We obtain the prediction on the training data from:

$$\hat{Y} \quad = \quad \Phi B \tag{15}$$

$$= \quad \Phi \Phi^T U (T^T \Phi \Phi^T U)^{-1} T^T Y \tag{16}$$

$$= \quad K U (T^T K U)^{-1} T^T Y. \tag{17}$$

To exclude the bias term, we assume that the responses and the predictors are set to have zero mean in the feature space by applying the following procedure to test kernel K_t and training kernel K [31]:

$$K_t \quad \leftarrow \quad \left(K_t - \frac{1}{n} \mathbf{1}_{n_t} \mathbf{1}_n^T K \right) \left(I_n - \frac{1}{n} \mathbf{1}_n \mathbf{1}_n^T \right) \tag{18}$$

$$K \quad \leftarrow \quad \left(I_n - \frac{1}{n} \mathbf{1}_n \mathbf{1}_n^T \right) K \left(I_n - \frac{1}{n} \mathbf{1}_n \mathbf{1}_n^T \right), \tag{19}$$

where $\mathbf{1}_n$ represents the n-length vector whose n elements are 1. Note that n and n_t represent the number of training and test samples, respectively.

In the following section of this paper, we investigate three kernel functions: 1) a second order polynomial kernel $k(x, x') = (x^T x' + 1)^2$, referred to as KPLS-Poly(2), 2) a third order polynomial kernel $k(x, x') = (x^T x' + 1)^3$, referred to as KPLS-Poly(3), 3) a Gaussian kernel $k(x, x') = \exp(-\gamma \|x - x'\|)^2)$, referred to as KPLS-Gauss, where γ is a hyper parameter and set to the inverse of the median of the Euclidian distance of data points.

Classification

In addition to predicting clinical measures, our aim is to classify subjects into depressed patients and healthy controls using the predicted value of clinical measures for objective diagnosis. We evaluate generalization of binary classifiers using linear discriminant analysis (LDA). Given the training data $\mathbb{D}_{tr} = \{\mathbf{x}_{tr}, \mathbf{y}_{tr}, \mathbf{z}_{tr}\}$ and test data $\mathbb{D}_{te} = \{\mathbf{x}_{te}, \mathbf{y}_{te}, \mathbf{z}_{te}\}$, $\mathbf{x} \in \mathbb{R}^p$, $\mathbf{y} \in \mathbb{R}^q$, and $\mathbf{z} \in \{0, 1\}$ represent functional connectivity as predictors, clinical measures as responses, and binary labels (i.e. 0 is patients and 1 is healthy controls), respectively. In the prediction phase, our objective is to learn the function $f_B : \mathbb{R}^p \to \mathbb{R}^q$, which, given predictors, $\mathbf{x}_{tr}$, and responses, $\mathbf{y}_{tr}$, assigns predictors to the most probable values of $\mathbf{y}$. The prediction on the training dataset is $\hat{\mathbf{y}}_{tr} = f_B(\mathbf{x}_{tr})$. In the next classification phase, our objective is to learn the classifier $g_w : \mathbb{R}^q \to \{0, 1\}$, which, given predicted responses, $\hat{\mathbf{y}}_{tr}$, and binary labels, $\mathbf{z}_{tr}$, assigns predicted responses to the most probable labels. Assigned labels on the test dataset are obtained as $\hat{\mathbf{z}}_{te} = g_w(\hat{\mathbf{y}}_{te}) = g_w(f_B(\hat{\mathbf{x}}_{te}))$. It is important to stress that the binary classifier is not trained on actual clinical measures, $\mathbf{y}_{tr}$, but on predicted values of $\hat{\mathbf{y}}_{tr}$.

In a previous study [13], the authors only identified the binary classifier $g'_w : \mathbb{R}^p \to 0, 1$, which, given functional connectivity, $\mathbf{x}_{tr}$, and binary labels, $\mathbf{z}_{tr}$, assigns functional connectivity directly to binary labels. By exploiting the predicted result of clinical measures, it may be possible to improve classification performance. We compared two scenarios, i.e. i) classification of patients and healthy controls using LDA from predicted clinical measures with KPLS (with KPLS-Gauss, KPLS-Poly(3), and KPLS-Poly(2)), PLS, and ordinary least squares regression (OLS), ii) classification of patients and healthy controls by means of LDA and SVM from functional connectivity directly. Note that we perform feature selection before scenario 2) by calculating connection-wise t-tests to determine the connections with different group means,

represented by t-scores. We select the *M* functional connections with the highest absolute t-scores. *M* is optimized by cross validations.

Pre-screening

Even though PLS can cope with high-dimensional, co-linear datasets, we prescreened variables depending on their relevance to responses in the following way.

Based on Pearson correlation coefficients, ρ_{rl}, between the *r*-th functional connection and the *l*-th clinical measures, we define the empirical relevance of the *r*-th functional connection as

$$R_r = \sum_{l=1}^{4} \rho_{rl}^2, \qquad r = 1, \ldots, p, \qquad (20)$$

where *p* is the total number of functional connections.

These functional connections are ranked according to their empirical relevance, $\{R_r\}_{r=1}^{p}$, and only *M* relevant functional connections are used in following procedure. The optimal number for M was determined through nested leave-one-out cross-validation.

Nested leave-one-out cross validation

Conventionally, cross validation is employed to assure generalization ability of a model or to evaluate optimal parameters. Since we have to account for both generalization ability and parameter optimization, we made use of nested leave-one-out cross validation (LOOCV), which consisted of outer and inner LOOCV. The outer LOOCV repeats iterations that split the whole set of samples into a single outer validation sample used to evaluate the generalization ability, and an outer training set for model estimation. The inner loop of LOOCV is performed on the outer training set to optimize two parameters, *M* and *L*, the number of selected predictor variables and the number of components, respectively. The pair of parameters that achieves the lowest root mean squared error based on the inner validation sample are adopted as optimal parameters and used to evaluate the model using the outer LOOCV. These steps are repeated until each sample has served as the validation sample.

Age

Age is significantly correlated with three clinical measures ([Table 2]). In general, age matching performed on different diagnostic groups reduces sample size, causing poor performance. To avoid this problem, we investigated three models, i.e. (i) a model with age as a response (denoted by output-age), (ii) a model with age as a predictor (denoted by input-age), and (iii) a model without age (denoted by no-age). By incorporating age into our model, we can cope with age differences among subjects and can fairly evaluate prediction performance.

Interpretation

Interpretation of each latent component projected from input and output data gives novel insights into the relationship between functional connectivity and clinical measures. In the framework of PLS, loading matrices, *P* and *C*, indicate contributions from predictor variables and response variables to each latent component (see Eqs ([10]) and ([11])). The (*i*, *j*)-element of the loading matrix, *P*, represents the contribution of the *i*-th functional connection to the *j*-th latent component. Similarly, the (*i*, *j*)-element of the loading matrix, *C*, represents the contribution of the *i*-th clinical measure to the *j*-th latent component. Note that due to subject variability, values of P_{ij} and C_{ij} vary depending on the training set used.
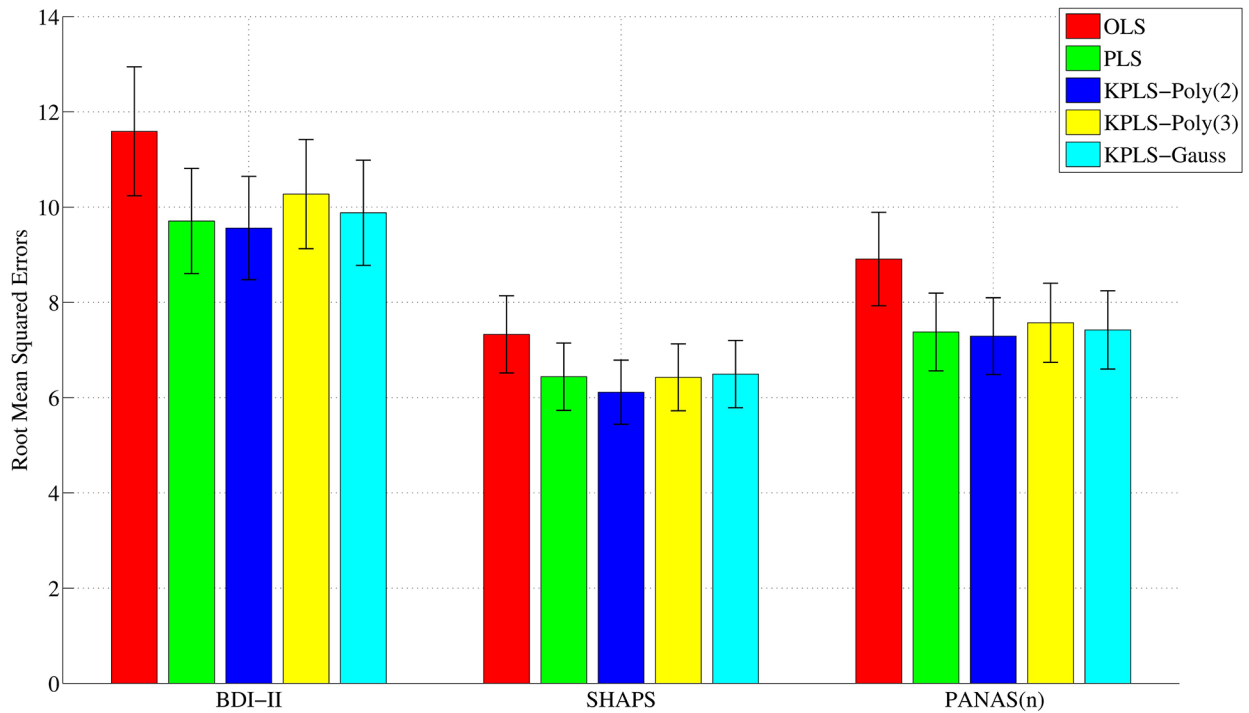
Fig 2. Comparison of predicted performance by means of the root mean squared errors. Linear and kernel variants of PLS achieved significantly better performance than did OLS in all clinical scores. Subject age was used as the output along with clinical scores (output-age model).

https://doi.org/10.1371/journal.pone.0179638.g002

Results

Regression performance

We compared the prediction performance of PLS, its kernel variants, and other methods by means of the root mean squared error (RMSE) of the predicted clinical measures in nested leave-one-out cross validation (see Methods). Kernel PLS with a second-order polynomial kernel (KPLS-Poly(2)) achieved the lowest RMSE (9.56 for BDI-II, 6.11 for SHAPS, and 7.29 for PANAS(n)) (Fig 2). This performance was significantly better than that of ordinary least squares regression (OLS) (11.6 for BDI-II, 7.33 for SHAPS, and 8.91 for PANAS(n)) and comparable to that of other variants of PLS applied in our study, suggesting that projection of data into a low-dimensional space was beneficial to regression performance. All statistical comparisons were adjusted for multiplicity using the Bonferroni-Holm method with significance level, $\alpha = 0.05$.

Next, to evaluate the best way of incorporating age into our regression models, we compared RMSE of the output-age, input-age, and no-age models. In our study, incorporating age into our regression model as a response (output-age) achieved significantly better performance than that of the input-age and no-age models (Fig 3). The details were listed in Supporting Information (S1, S2 and S3 Tables). All statistical comparisons were adjusted for multiplicity using the Bonferroni-Holm method with significance level, $\alpha = 0.05$.

The correlation coefficient of actual and predicted values for BDI-II, SHAPS, and PANAS (n) in the case of KPLS-Poly(2) were $r = 0.541, 0.591, 0.563$, respectively. The relationship between predicted and actual values of BDI-II for KPLS-Poly(2) was exemplified (Fig 4). This
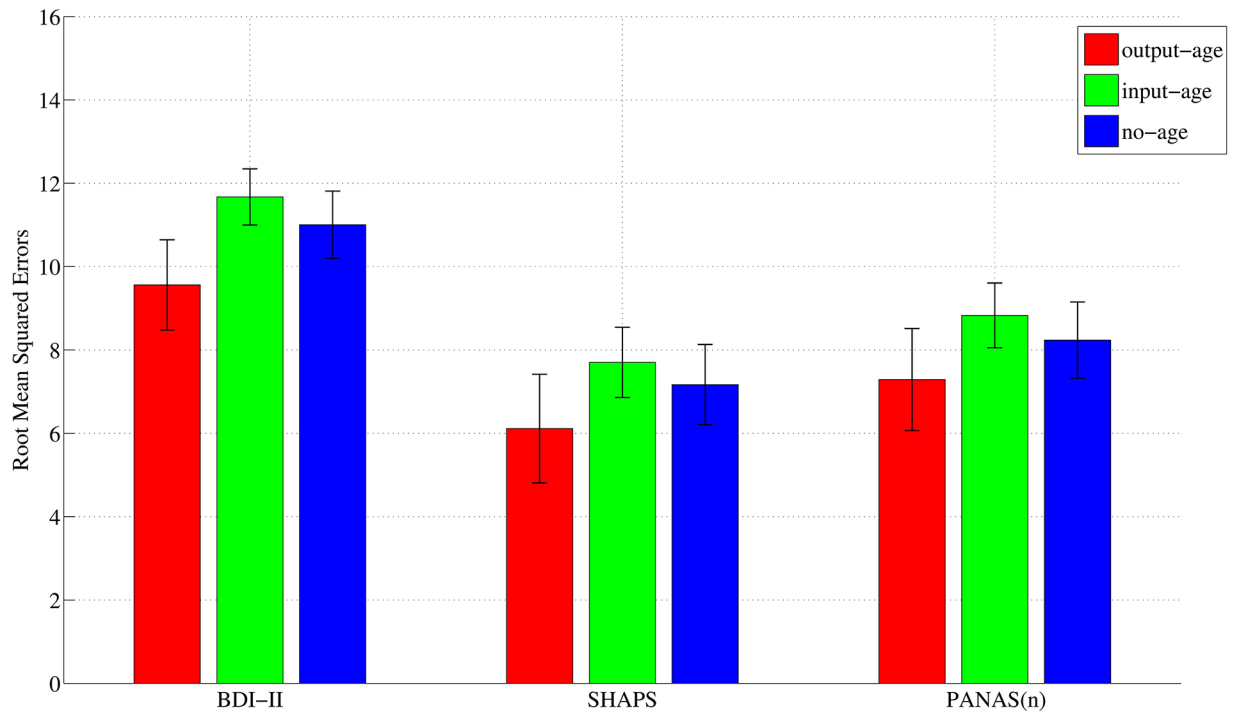
Fig 3. Root mean squared errors in KPLS-Poly(2). KPLS-Poly(2) achieved significantly better performance in output-age model than in other models.
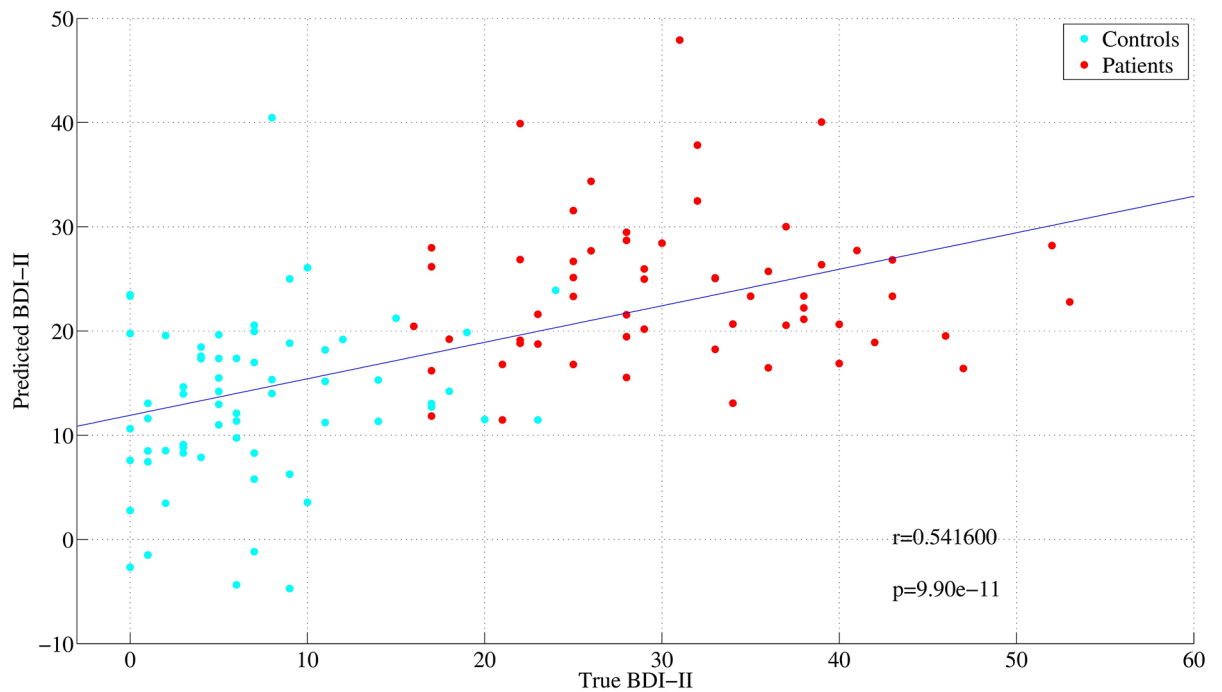
Fig 4. Actual and predicted values of BDI-II. BDI-II were well predicted by KPLS-Poly(2) with RMSE = 9.56 and r = 0.541 ($p < 10^{-10}$). Red and blue points represent patients and healthy controls, respectively.
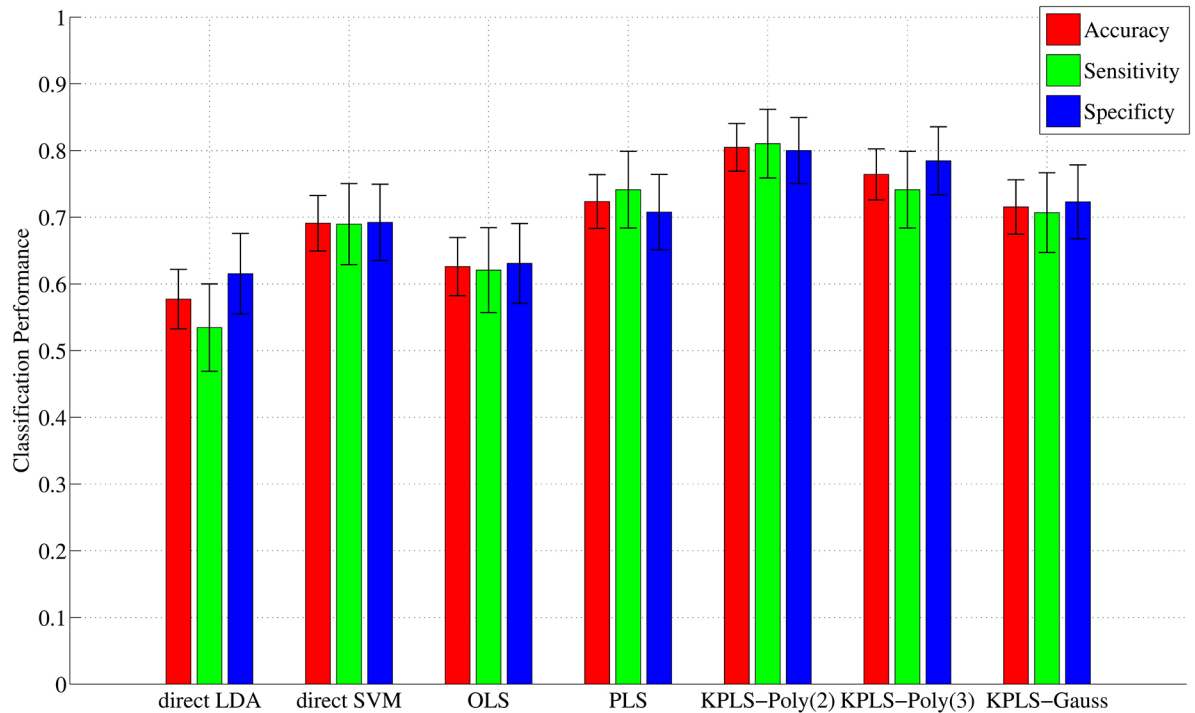
Fig 5. Classification accuracy, sensitivity, and specificity. KPLS-Poly(2) followed by LDA achieves the best performance (accuracy = 80.5%, sensitivity = 81.0%, and specificity = 80.0%).

result was comparable to that of Zhang et al. (2011) [14]; however, the number of subjects in our study was larger than in theirs, reconfirming validity of the results.

The optimal number of retained features M^* identified by pre-screening and using the latent component L^* identified with nested LOOCV were 40 and 3, respectively, suggesting that reduction of feature size was relevant for improvement of PLS prediction accuracy.

Classification performance

Projecting the original data onto a low-dimensional space was expected to improve classification accuracy. To verify the benefit of projection, several classification methods were performed and evaluated using accuracy, sensitivity, and specificity (Fig 5). The details were listed in Supporting Information (S4 and S5 Tables). In our study, KPLS-Poly(2) followed by LDA achieved the best accuracy 80.5% (sensitivity 81.0% and specificity 80.0%), which is significantly better than the 57.7% accuracy of direct LDA (sensitivity 53.4%, and specificity 61.5%) and 69.1% accuracy of direct SVM (sensitivity 69.0%, and specificity 69.2%). This result indicates that it was beneficial to exploit the prediction model for clinical measures in order to build a classification model. In addition, KPLS-Poly(2) followed by LDA also achieved significantly better accuracy than the 62.6% accuracy of OLS followed by LDA (sensitivity 62.1% and specificity 63.1%), indicating that considering a latent space in a regression model was beneficial to final classification. Accuracy did not differ significantly between PLS and kernel variants. All statistical tests were based on approximation with the normal and adjusted for multiplicity using the Bonferroni-Holm method with significance level $\alpha = 0.05$.
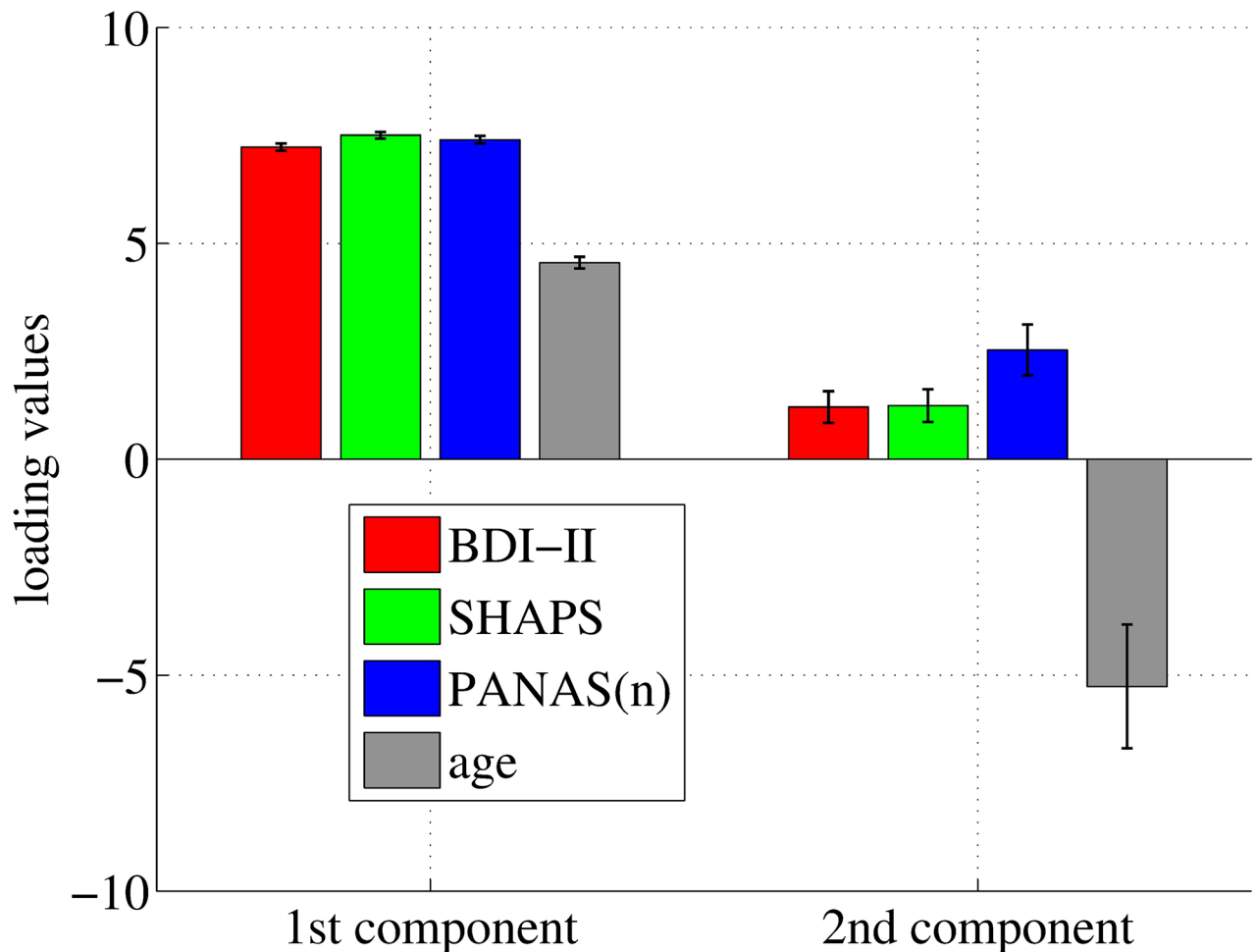
Fig 6. Loading matrix *C*. The matrix indicates contributions of each clinical measure to the first two latent components. Age has a negative influence on the second latent component. Error bars indicate the standard deviation.

Interpretation

In our study, three clinical scores showed almost equally positive influences on the first component, and age also had a positive influence as well. However, age showed a strong negative influence on the second component, in contrast to the clinical scores (Fig 6).

Latent space representation of subjects showed that the first component explained most depression severity in comparison with the second component (Fig 7). This is consistent with the results of loading matrix *C*. Note that since the optimal number of latent components, in terms of minimizing regression error, was 3, the second and the third components are thought to contain some information about scores.

In order to validate the effect of age, especially in the second component, all subjects were grouped into young (age 20–31, 41 subjects), middle (age 31–43, 41 subjects), and old (age 44–73, 41 subjects) groups. Note we simply divided the subjects in three equal-sized groups for convenience, "young", "middle", and "old". They are relative, not absolute age classes. Latent variables of old subjects in the second component were significantly lower than those of young
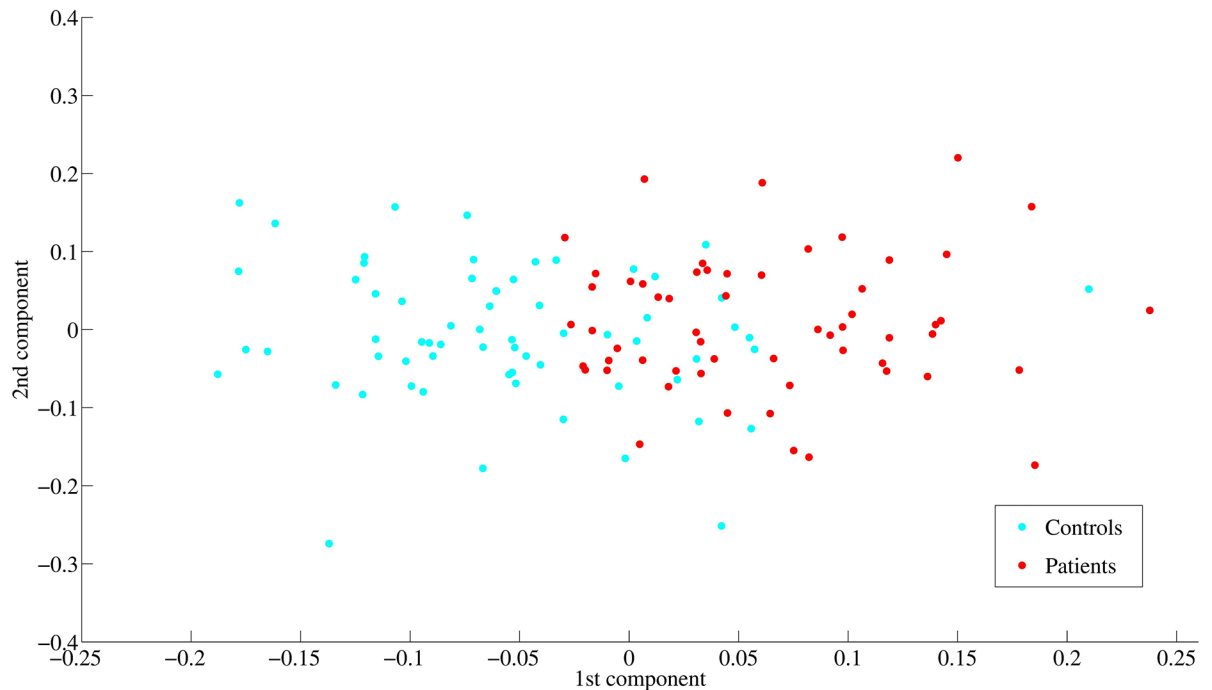
Fig 7. Scatter plot of the latent variables in the first two latent components generated from KPLS-Poly(2). Red and blue dots represent patients and healthy controls, respectively. The two groups are separated mainly by the first latent component.

https://doi.org/10.1371/journal.pone.0179638.g007

and middle subjects ($p < 10^{-5}$ by Wilcoxon Rank-Sum Test), suggesting that old patients have distinctive patterns in the second latent space [5] (Fig 8).

Evaluation of loading matrix, P, reveals functional connections relevant to each latent component. Especially, the first column of P, corresponding to the first component responsible for discrimination of each diagnostic group, is expected to reveal useful insights about the effect of functional connections on depression symptoms. Even though the performance of KPLS-Poly (2) in prediction and classification was comparable to or better than that of linear PLS, patterns of significant loadings were consistent in our experiments. For reasons of interpretation, we therefore focus on the loading matrix of the linear terms in following sections.

BrainNet Viewer [32] (http://www.nitrc.org/projects/bnv/) was used to visualize the top 10 connections with positive and negative loadings for the first component (Figs 9 and 10). In this figure, many regions involved in the default mode network (DMN), as well as *the left supplementary motor area*, *the right superior frontal gyrus*, and *the insula*, were relevant. In addition, some functional connectivity between *the right cuneus* and regions involved in *the cerebellum* were negatively correlated with the first component.

Discussion

MacIntosh et al. (1996) first introduced PLS analysis into the field of neuroimaging in order to extract common information between brain activity and exogenous information, such as experimental or behavioral data [3, 4]. In particular, behavioral data are increasingly used to extract associated brain activity patterns for various types of psychological diseases, such as Alzheimer's disease [33], obsessive-compulsive disorder [34], and schizophrenia [35]. In these studies, neuropsychological test scores are used as behavioral data, in addition to the labels
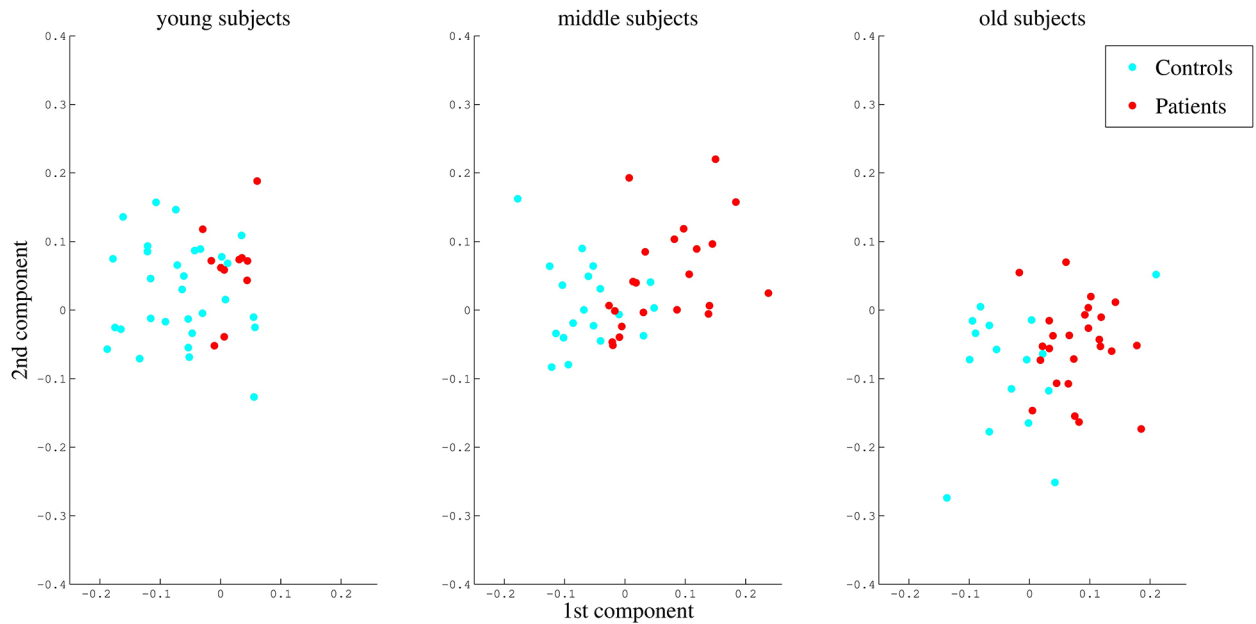
Fig 8. Scatter plot of subjects separated on the basis of the two latent components generated from KPLS-Poly(2) for young, middle, and old subjects. Old subjects have significantly lower values in the second component ($p < 10^{-5}$ by Wilcoxon Rank-Sum Test).

that represent diagnostic groups and age. To the best of our knowledge, this is the first study to investigate associations between functional connectivity in the whole brain and multiple clinical measures for depressed patients, using PLS and its kernel variants.

Diagnosis based on resting-state functional connectivity is a challenging task due to the high-dimensionality and co-linearity of data. Recent studies have demonstrated that depressed patients can be distinguished from healthy controls by means of their functional connectivity by applying conventional methods, such as support vector machine [9, 13]. Since binary labels are ultimately abstracted information about depression that ignores the severity of symptoms, it is worth considering more detailed information, such as BDI-II, SHAPS, and PANAS(n) to build more sophisticated models. Our study demonstrated that projecting functional connectivity data into a low-dimensional latent space, can predict clinical measures, and can also improve depression diagnostic accuracy.

To separately identify neural circuits associated with anhedonia and negative mood is a challenging task. A psychopathological study suggests that these primary symptoms result from different neural circuits and from alternation of different neurotransmitters [16]. Our results show that SHAPS and PANAS(n) are highly correlated and contributed quite similarly to each latent component (Fig 6), suggesting that further investigation and different approaches may be required to support psychopathological studies from the point of data driven analysis.

We also evaluated extended AAL generated by subdividing standard AAL into 600 regions to examine if the finer atlas could be used to improve the prediction of the clinical scores [36]. The performance was significantly worse than that of standard AAL (S6 Table) and the selected functional connections were inconsistent. Since analysis of brain imaging data with limited sample size highly depends on the choice of ROI, the finer atlas does not necessarily
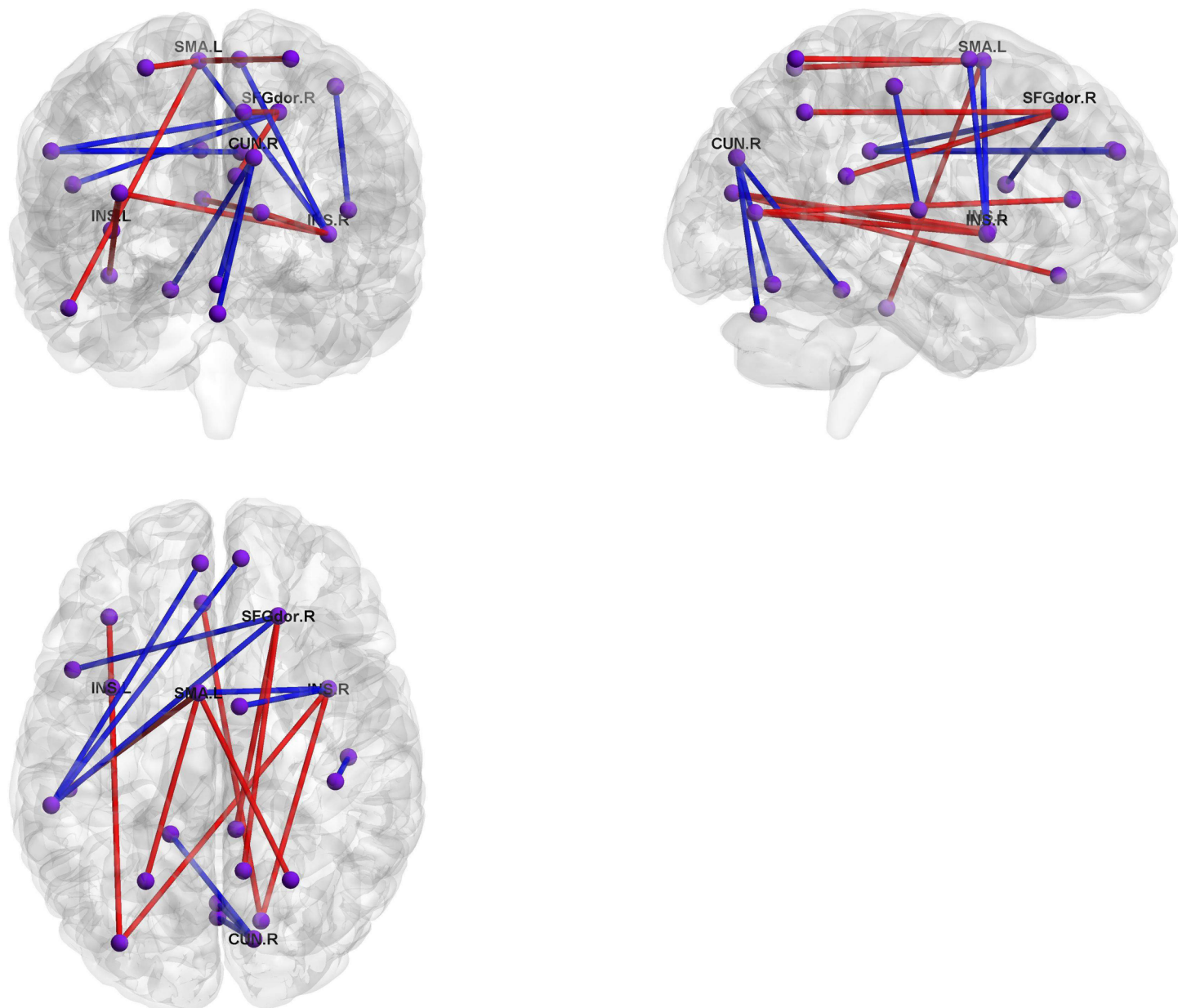
Fig 9. Contributing functional connectivity in first latent component. Red and blue lines represent positive and negative loadings, respectively. SFGdor. R: *right superior frontal gyrus*, INS: *insula*, SMA.L: *left supplementary motor area*, CUN.R:*right cuneus*.

provide better prediction performance. Therefore, it is fare to note that further research is required to validate the best atlas.

Contributing brain regions

Identification of relevant brain regions in functional connectivity analysis yielded the following three observations: (1) connections between the default mode network and other regions, such as *the right superior frontal gyrus* and *the left supplementary motor area* are relevant (2) *the*
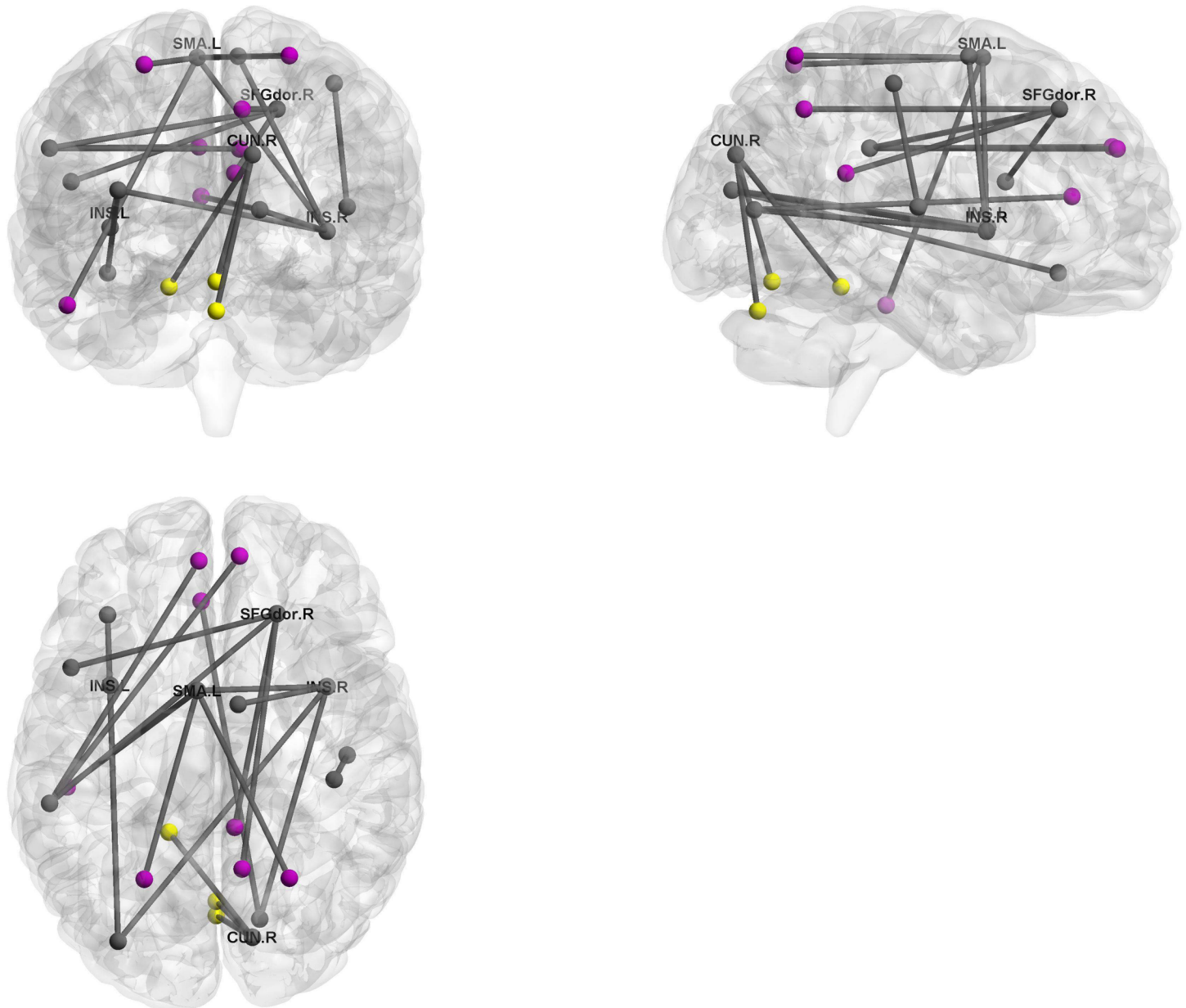
Fig 10. Contributing functional connectivity in first latent component. Purple and yellow nodes represent brain areas within the default mode network and *cerebellum*, respectively. SFGdor.R: *right superior frontal gyrus*, INS: *insula*, SMA.L: *left supplementary motor area*, CUN.R: *right cuneus*.

left and right insulas in both hemispheres are relevant, (3) connections between *the cerebellum* and *the right cuneus* are relevant.

First, the default mode network (DMN) shows synchronized deactivation during cognitive tasks and is thought to be related to major depressive disorder [37–40]. Our study supports these results, indicating that many relevant connections are related to the DMN, such as *the right posterior cingulum*, *the right precuneus*, and *the superior parietal gyrus*. The DMN contributes positive connections with *the right superior frontal gyrus* and *the left supplementary motor*

area. The superior frontal gyrus, as a critical region in cognitive tasks, was previously reported to be associated with depression [41]. While *the supplementary motor area* is known to be responsible for motor control, it is also reportedly related to some subtype of depression [42]. Our results support these results.

Second, our results suggest that *the insula* is associated with depression. Some meta-analysis of PET and fMRI studies revealed that *the insula* plays an important role in regulation of emotion [43, 44]. Similarly, resting-state fMRI studies have indicated that *the insula* is directly associated with depression [45, 46].

Finally, we showed that three connections between *the right cuneus*, located in the visual cortical area, and *the cerebellum*, negatively influence depression. While visual processing is believed not to be affected in depression, some previous studies have suggested that it is associated with bipolar disorder [47]. Moreover, regional homogeneity (ReHo) interpreted as a measure of localized synchrony in resting-state fMRI was decreased [48]. *The cerebellum* is usually considered to be responsible for motion control, but our results indicate that it may also be involved in regulation of mood and cognitive processing associated with symptoms of depression. Some fMRI studies demonstrate that this area is responsible for various types of information processing [49, 50], and resting-state functional connectivity studies imply that *the cerebellum* may be critical for the distinction between depressed patients and healthy controls [13, 51].

Conclusion

In summary, we employed partial least squares regression and its kernel variants to predict clinical measures of subjects using resting-state functional connectivity. Diagnosis of depression based on predicted clinical scores performed better than classification algorithms attempting diagnoses directly from functional connectivity. Moreover, analysis of latent variables identified functional networks relevant to the diagnosis of depression. These results suggest that a low-dimensional representation derived using PLS is beneficial for objective diagnosis. Further investigations are required to separate the two neural circuits associated with two primary symptoms, anhedonia and negative mood.

Supporting information

S1 Appendix. Partial least squares regression.
(PDF)

S1 Table. Root mean squared errors in output-age. Bold figures represent the best achievement.
(PDF)

S2 Table. Root mean squared errors in input-age.
(PDF)

S3 Table. Root mean squared errors in no-age.
(PDF)

S4 Table. Classification performance. KPLS-Poly(2) followed by LDA significantly outperformed direct LDA, SVM, and OLS followed by LDA in accuracy (adjusted for multiplicity using the Bonferroni-Holm method with significance level $\alpha = 0.05$).
(PDF)

S5 Table. Classification performance without subjects over 60. KPLS-Poly(2) followed by LDA significantly outperformed direct LDA and OLS followed by LDA in accuracy (adjusted

for multiplicity using the Bonferroni-Holm method with significance level $\alpha = 0.05$).
(PDF)

S6 Table. Root mean squared errors in extended AAL. Performance with the standard AAL outperformed extended AAL significantly. (adjusted for multiplicity using the Bonferroni-Holm method with significance level $\alpha = 0.05$).
(PDF)

Acknowledgments

Author Contributions

Conceptualization: KY YS JY KD.

Data curation: MT GO YO SY.

Formal analysis: KY.

Funding acquisition: SY KD.

Investigation: KY.

Methodology: KY.

Project administration: KD.

Resources: MT GO YO SY.

Software: KY.

Supervision: KD.

Validation: KY.

Visualization: KY.

Writing – original draft: KY.

Writing – review & editing: YS JY MT GO YO SY.

References

1. Shimizu Y, Yoshimoto J, Toki S, Takamura M, Yoshimura S, Okamoto Y, et al. Toward probabilistic diagnosis and understanding of depression based on functional MRI data analysis with logistic group LASSO. PLoS ONE. 2015; 10(5):e0123524. https://doi.org/10.1371/journal.pone.0123524 PMID: 25932629

2. Wold H, et al. Soft modeling by latent variables: the nonlinear iterative partial least squares approach. Perspectives in Probability and Statistics, papers in honour of MS Bartlett. 1975; p. 520–540.

3. McIntosh A, Bookstein F, Haxby JV, Grady C. Spatial pattern analysis of functional brain images using partial least squares. NeuroImage. 1996; 3(3):143–157. https://doi.org/10.1006/nimg.1996.0016 PMID: 9345485

4. McIntosh AR, Lobaugh NJ. Partial least squares analysis of neuroimaging data: applications and advances. NeuroImage. 2004; 23:S250–S263. https://doi.org/10.1016/j.neuroimage.2004.07.020 PMID: 15501095

5. Chen K, Reiman EM, Huan Z, Caselli RJ, Bandy D, Ayutyanont N, et al. Linking functional and structural brain images with multivariate network analyses: a novel application of the partial least square method.

NeuroImage. 2009; 47(2):602–610. https://doi.org/10.1016/j.neuroimage.2009.04.053 PMID: 19393744

6.   Ziegler G, Dahnke R, Winkler AD, Gaser C. Partial least squares correlation of multivariate cognitive abilities and local brain structure in children and adolescents. NeuroImage. 2013; 82:284–294. https://doi.org/10.1016/j.neuroimage.2013.05.088 PMID: 23727321

7.   Giessing C, Fink GR, Rösler F, Thiel CM. fMRI data predict individual differences of behavioral effects of nicotine: a partial least square analysis. Journal of Cognitive Neuroscience. 2007; 19(4):658–670. https://doi.org/10.1162/jocn.2007.19.4.658 PMID: 17381256

8.   Yahata N, Morimoto J, Hashimoto R, Lisi G, Shibata K, Kawakubo Y, et al. A small number of abnormal brain connections predicts adult autism spectrum disorder. Nature Communications. 2016; 7. https://doi.org/10.1038/ncomms11254 PMID: 27075704

9.   Craddock RC, Holtzheimer PE, Hu XP, Mayberg HS. Disease state prediction from resting state functional connectivity. Magnetic Resonance in Medicine. 2009; 62(6):1619–1628. https://doi.org/10.1002/mrm.22159 PMID: 19859933

10.   Greicius MD, Flores BH, Menon V, Glover GH, Solvason HB, Kenna H, et al. Resting-state functional connectivity in major depression: abnormally increased contributions from subgenual cingulate cortex and thalamus. Biological Psychiatry. 2007; 62(5):429–437. https://doi.org/10.1016/j.biopsych.2006.09.020 PMID: 17210143

11.   Veer IM, Beckmann CF, Van Tol MJ, Ferrarini L, Milles J, Veltman DJ, et al. Whole brain resting-state analysis reveals decreased functional connectivity in major depression. Frontiers in Systems Neuroscience. 2010; 4. https://doi.org/10.3389/fnsys.2010.00041 PMID: 20941370

12.   Zhang J, Wang J, Wu Q, Kuang W, Huang X, He Y, et al. Disrupted brain connectivity networks in drug-naive, first-episode major depressive disorder. Biological Psychiatry. 2011; 70(4):334–342. https://doi.org/10.1016/j.biopsych.2011.05.018 PMID: 21791259

13.   Zeng LL, Shen H, Liu L, Wang L, Li B, Fang P, et al. Identifying major depression using whole-brain functional connectivity: a multivariate pattern analysis. Brain. 2012; 135(5):1498–1507. https://doi.org/10.1093/brain/aws059 PMID: 22418737

14.   Zhang X, Yaseen ZS, Galynker II, Hirsch J, Winston A. Can depression be diagnosed by response to mother's face? A personalized attachment-based paradigm for diagnostic fMRI. PLoS ONE. 2011; 6 (12):e27253. https://doi.org/10.1371/journal.pone.0027253 PMID: 22180777

15.   Beck AT, Steer RA, Ball R, Ranieri WF. Comparison of Beck Depression Inventories-IA and -II in psychiatric outpatients. Journal of Personality Assessment. 1996; 67(3):588–597. https://doi.org/10.1207/s15327752jpa6703_13 PMID: 8991972

16.   Hasler G, Northoff G. Discovering imaging endophenotypes for major depression. Molecular Psychiatry. 2011; 16(6):604–619. https://doi.org/10.1038/mp.2011.23 PMID: 21602829

17.   Snaith R, Hamilton M, Morley S, Humayan A, Hargreaves D, Trigwell P. A scale for the assessment of hedonic tone the Snaith-Hamilton Pleasure Scale. The British Journal of Psychiatry. 1995; 167(1):99–103. https://doi.org/10.1192/bjp.167.1.99 PMID: 7551619

18.   Watson D, Clark LA, Tellegen A. Development and validation of brief measures of positive and negative affect: the PANAS scales. Journal of Personality and Social Psychology. 1988; 54(6):1063. https://doi.org/10.1037/0022-3514.54.6.1063 PMID: 3397865

19.   Katona CL, Watkin V. Depression in old age. Reviews in Clinical Gerontology. 1995; 5(04):427–441. https://doi.org/10.1017/S095925980000486X

20.   Yoshida K, Shimizu Y, Yoshimoto J, Toki S, Okada G, Takamura M, et al. Resting state functional connectivity explains individual scores of multiple clinical measures for major depression. In: Bioinformatics and Biomedicine (BIBM), 2015 IEEE International Conference on. IEEE; 2015. p. 1078–1083.

21.   Lecrubier Y, Sheehan DV, Weiller E, Amorim P, Bonora I, Harnett Sheehan K, et al. The Mini International Neuropsychiatric Interview (MINI). A short diagnostic structured interview: reliability and validity according to the CIDI. European Psychiatry. 1997; 12(5):224–231. https://doi.org/10.1016/S0924-9338(97)83296-8

22.   American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, 4th Edition, Text Revision (DSM-IV-TR). American Psychiatric Association; 2000.

23.   Ehring T, Tuschen-Caffier B, Schnülle J, Fischer S, Gross JJ. Emotion regulation and vulnerability to depression: spontaneous versus instructed use of emotion suppression and reappraisal. Emotion. 2010; 10(4):563. https://doi.org/10.1037/a0019010 PMID: 20677873

24.   Van Dijk KR, Sabuncu MR, Buckner RL. The influence of head motion on intrinsic functional connectivity MRI. NeuroImage. 2012; 59(1):431–438. https://doi.org/10.1016/j.neuroimage.2011.07.044 PMID: 21810475

25. Zeng LL, Wang D, Fox MD, Sabuncu M, Hu D, Ge M, et al. Neurobiological basis of head motion in brain imaging. Proceedings of the National Academy of Sciences. 2014; 111(16):6058–6062. https://doi.org/10.1073/pnas.1317424111

26. Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, et al. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. NeuroImage. 2002; 15(1):273–289. https://doi.org/10.1006/nimg.2001.0978 PMID: 11771995

27. Manne R. Analysis of two partial-least-squares algorithms for multivariate calibration. Chemometrics and Intelligent Laboratory Systems. 1987; 2(1):187–197. https://doi.org/10.1016/0169-7439(87)80096-5

28. Rosipal R, Trejo LJ. Kernel partial least squares regression in reproducing kernel hilbert space. Journal of Machine Learning Research. 2002; 2:97–123.

29. Rännar S, Lindgren F, Geladi P, Wold S. A PLS kernel algorithm for data sets with many variables and fewer objects. Part 1: Theory and algorithm. Journal of Chemometrics. 1994; 8(2):111–125. https://doi.org/10.1002/cem.1180080204

30. Bennett K, Embrechts M. An optimization perspective on kernel partial least squares regression. NATO Science Series sub series III Computer and Systems Sciences. 2003; 190:227–250.

31. Schölkopf B, Smola A, Müller KR. Nonlinear component analysis as a kernel eigenvalue problem. Neural Computation. 1998; 10(5):1299–1319. https://doi.org/10.1162/089976698300017467

32. Xia M, Wang J, He Y. BrainNet Viewer: a network visualization tool for human brain connectomics. PLoS ONE. 2013; 8(7):e68910. https://doi.org/10.1371/journal.pone.0068910 PMID: 23861951

33. Price J, Ziolko S, Weissfeld L, Klunk W, Lu X, Hoge J, et al. Quantitative and statistical analyses of PET imaging studies of amyloid deposition in humans. In: Nuclear Science Symposium Conference Record, 2004 IEEE. vol. 5. IEEE; 2004. p. 3161–3164.

34. Menzies L, Achard S, Chamberlain SR, Fineberg N, Chen CH, Del Campo N, et al. Neurocognitive endophenotypes of obsessive-compulsive disorder. Brain. 2007; 130(12):3223–3236. https://doi.org/10.1093/brain/awm205 PMID: 17855376

35. Nestor PG, O'Donnell BF, McCarley RW, Niznikiewicz M, Barnard J, Shen ZJ, et al. A new statistical method for testing hypotheses of neuropsychological/MRI relationships in schizophrenia: partial least squares analysis. Schizophrenia Research. 2002; 53(1):57–66. https://doi.org/10.1016/S0920-9964(00)00171-7 PMID: 11728838

36. Hermundstad AM, Bassett DS, Brown KS, Aminoff EM, Clewett D, Freeman S, et al. Structural foundations of resting-state and task-based functional connectivity in the human brain. Proceedings of the National Academy of Sciences. 2013; 110(15):6169–6174. https://doi.org/10.1073/pnas.1219562110

37. Zhu X, Wang X, Xiao J, Liao J, Zhong M, Wang W, et al. Evidence of a dissociation pattern in resting-state default mode network connectivity in first-episode, treatment-naive major depression patients. Biological Psychiatry. 2012; 71(7):611–617. https://doi.org/10.1016/j.biopsych.2011.10.035 PMID: 22177602

38. Zeng LL, Shen H, Liu L, Hu D. Unsupervised classification of major depression using functional connectivity MRI. Human Brain Mapping. 2014; 35(4):1630–1641. https://doi.org/10.1002/hbm.22278 PMID: 23616377

39. Kaiser RH, Andrews-Hanna JR, Wager TD, Pizzagalli DA. Large-scale network dysfunction in major depressive disorder: a meta-analysis of resting-state functional connectivity. JAMA psychiatry. 2015; 72(6):603–611. https://doi.org/10.1001/jamapsychiatry.2015.0071 PMID: 25785575

40. Mulders PC, van Eijndhoven PF, Schene AH, Beckmann CF, Tendolkar I. Resting-state functional connectivity in major depressive disorder: A review. Neuroscience and Biobehavioral Reviews. 2015; 56:330–344. https://doi.org/10.1016/j.neubiorev.2015.07.014 PMID: 26234819

41. Fitzgerald PB, Laird AR, Maller J, Daskalakis ZJ. A meta-analytic study of changes in brain activation in depression. Human Brain Mapping. 2008; 29(6):683–695. https://doi.org/10.1002/hbm.20426 PMID: 17598168

42. Exner C, Lange C, Irle E. Impaired implicit learning and reduced pre-supplementary motor cortex size in early-onset major depression with melancholic features. Journal of Affective Disorders. 2009; 119(1-3):156–62. https://doi.org/10.1016/j.jad.2009.03.015 PMID: 19345999

43. Phan KL, Wager T, Taylor SF, Liberzon I. Functional neuroanatomy of emotion: a meta-analysis of emotion activation studies in PET and fMRI. NeuroImage. 2002; 16(2):331–348. https://doi.org/10.1006/nimg.2002.1087 PMID: 12030820

44. Hamilton JP, Etkin A, Furman DJ, Lemus MG, Johnson RF, Gotlib IH. Functional neuroimaging of major depressive disorder: a meta-analysis and new integration of baseline activation and neural

response data. American Journal of Psychiatry. 2012;. https://doi.org/10.1176/appi.ajp.2012.11071105 PMID: 22535198

45. Yao Z, Wang L, Lu Q, Liu H, Teng G. Regional homogeneity in depression and its relationship with separate depressive symptom clusters: a resting-state fMRI study. Journal of Affective Disorders. 2009; 115(3):430–438. https://doi.org/10.1016/j.jad.2008.10.013 PMID: 19007997

46. Liu Z, Xu C, Xu Y, Wang Y, Zhao B, Lv Y, et al. Decreased regional homogeneity in insula and cerebellum: a resting-state fMRI study in patients with major depression and subjects at high risk for major depression. Psychiatry Research: Neuroimaging. 2010; 182(3):211–215. https://doi.org/10.1016/j. pscychresns.2010.03.004 PMID: 20493670

47. Haldane M, Cunningham G, Androutsos C, Frangou S. Structural brain correlates of response inhibition in Bipolar Disorder I. Journal of Psychopharmacology. 2008;. https://doi.org/10.1177/ 0269881107082955 PMID: 18308812

48. Iwabuchi SJ, Krishnadas R, Li C, Auer DP, Radua J, Palaniyappan L. Localized connectivity in depression: a meta-analysis of resting state functional imaging studies. Neuroscience and Biobehavioral Reviews. 2015; 51:77–86. https://doi.org/10.1016/j.neubiorev.2015.01.006 PMID: 25597656

49. O'Reilly JX, Beckmann CF, Tomassini V, Ramnani N, Johansen-Berg H. Distinct and overlapping functional zones in the cerebellum defined by resting state functional connectivity. Cerebral Cortex. 2010; 20(4):953–965. https://doi.org/10.1093/cercor/bhp157 PMID: 19684249

50. Moulton EA, Elman I, Pendse G, Schmahmann J, Becerra L, Borsook D. Aversion-related circuitry in the cerebellum: responses to noxious heat and unpleasant images. Journal of Neuroscience. 2011; 31 (10):3795–3804. https://doi.org/10.1523/JNEUROSCI.6709-10.2011 PMID: 21389234

51. Liu L, Zeng LL, Li Y, Ma Q, Li B, Shen H, et al. Altered cerebellar functional connectivity with intrinsic connectivity networks in adults with major depressive disorder. PLoS ONE. 2012; 7(6):e39516. https:// doi.org/10.1371/journal.pone.0039516 PMID: 22724025