



OPEN

# Resting-state functional connectivity disruption between the left and right pallidum as a biomarker for subthreshold depression

Yosuke Sato<sup>1</sup>, Go Okada<sup>1</sup>, Satoshi Yokoyama<sup>1</sup>, Naho Ichikawa<sup>1,2</sup>, Masahiro Takamura<sup>3,4</sup>, Yuki Mitsuyama<sup>1</sup>, Ayaka Shimizu<sup>1</sup>, Eri Itai<sup>1</sup>, Hotaka Shinzato<sup>1</sup>, Mitsuo Kawato<sup>5</sup>, Noriaki Yahata<sup>6</sup> & Yasumasa Okamoto<sup>1</sup>✉

Although the identification of late adolescents with subthreshold depression (StD) may provide a basis for developing effective interventions that could lead to a reduction in the prevalence of StD and prevent the development of major depressive disorder, knowledge about the neural basis of StD remains limited. The purpose of this study was to develop a generalizable classifier for StD and to shed light on the underlying neural mechanisms of StD in late adolescents. Resting-state functional magnetic resonance imaging data of 91 individuals (30 StD subjects, 61 healthy controls) were included to build an StD classifier, and eight functional connections were selected by using the combination of two machine learning algorithms. We applied this biomarker to an independent cohort ( $n=43$ ) and confirmed that it showed generalization performance (area under the curve = 0.84/0.75 for the training/test datasets). Moreover, the most important functional connection was between the left and right pallidum, which may be related to clinically important dysfunctions in subjects with StD such as anhedonia and hyposensitivity to rewards. Investigation of whether modulation of the identified functional connections can be an effective treatment for StD may be an important topic of future research.

Recently, there has been growing interest in subthreshold depression (StD)<sup>1</sup>, which is defined as clinically relevant depressive symptoms not meeting the criteria for full-blown major depressive disorder (MDD)<sup>2</sup>. StD is highly prevalent among university students worldwide<sup>3</sup> and its prevalence seems to be increasing<sup>4</sup>. Late adolescent university students are reportedly vulnerable to depression and anxiety<sup>5-7</sup>, have a higher rate of depression compared to the general population<sup>8</sup>, and their depressive symptoms may be related to an increased risk for suicidal ideation and suicide attempts<sup>9</sup>. StD is associated with severe functional impairments and has a negative impact on the quality of life. In addition, StD is a risk factor for developing MDD<sup>6,10</sup>. Hence, the identification of late adolescents with StD may provide a basis for developing effective interventions that could lead to a reduction in the prevalence of StD and prevent the development of MDD. However, knowledge about the neural basis of StD remains limited, which makes it challenging to develop effective biomarkers and treatments<sup>11</sup>.

Some studies have evaluated impaired brain regions and abnormal patterns of brain connectivity to develop novel biomarkers for StD using resting-state functional magnetic resonance imaging (rs-fMRI)<sup>11-18</sup>. In early studies, it was reported that spontaneous neuronal activity, as measured by the amplitude of low-frequency fluctuations, is altered, that resting-state functional connections (FCs) are impaired, and that subcortical degree centrality is decreased and cortical degree centrality is increased in subjects with StD<sup>12,13,16</sup>. In studies pertaining

<sup>1</sup>Department of Psychiatry and Neurosciences, Hiroshima University, 1-2-3 Kasumi, Minami-ku, Hiroshima 734-8551, Japan. <sup>2</sup>Deloitte Analytics R&D, Deloitte Touche Tohmatsu LLC, Tokyo, Japan. <sup>3</sup>Department of Neurology, Shimane University, Matsue, Japan. <sup>4</sup>Center for Brain, Mind and KANSEI Research Sciences, Hiroshima University, Hiroshima, Japan. <sup>5</sup>ATR Brain Information Communication Research Laboratory Group, Kyoto, Japan. <sup>6</sup>Institute for Quantum Life Science, National Institutes for Quantum Science and Technology, Chiba, Japan. ✉email: oy@hiroshima-u.ac.jp

to brain networks, it was suggested that FCs between subregions of the anterior cingulate cortex are altered, that the FCs of the habenula within the default model network are increased, and that the FCs within salience networks are diminished in individuals with StD<sup>11,17</sup>. Recently, beyond the group-level analyses used in these previous studies, a data-driven approach using machine learning, which identifies phenotypes in a way that is clinically useful and can be applied to clinical diagnosis or prognosis, has gained attention<sup>19</sup>. In a recent investigation, by using the Support Vector Machine, which has been used frequently to identify imaging biomarkers in some diseases<sup>20</sup>, connectome-based biomarkers predicting StD associated with regions of the thalamus were identified, but with an insufficient generalization capability<sup>21</sup>. If the generalization capability to independent data is not sufficient, it is difficult to discuss the pathophysiology of the target disease through FC features selected by machine learning. Yahata et al.<sup>22</sup> developed an rs-fMRI-based classifier for autism spectrum disorder with generalization capability by combining two machine learning algorithms, L1-regularized sparse canonical correlation analysis (L1-SCCA)<sup>23</sup> and sparse logistic regression (SLR)<sup>24</sup>, to overcome the two major difficulties causing poor generalization performance, i.e., over-fitting and nuisance variables, and additionally by testing the classifier in an independent validation cohort<sup>22</sup>. By using the same method, we also developed a classifier for MDD with melancholic features that were generalized to an independent cohort and extracted critically important FCs<sup>25</sup>. To our knowledge, no study has attempted to identify an rs-fMRI-based biomarker for StD and confirmed its generalization capability. Therefore, the purpose of the present study was to develop a generalizable classifier for StD in late adolescents by using the combination of two machine learning algorithms, L1-SCCA and SLR<sup>22,25</sup>. Another purpose of this study was to shed light on the underlying neural mechanisms of StD by using the critically important extracted FCs. In late adolescents, the StD group has been characterized only by hyposensitivity to environmental rewards in previous studies, whereas the MDD group has been characterized by a higher frequency of avoidance and hyposensitivity to environmental rewards compared with the non-depression group<sup>26</sup>. Moreover, anhedonia is one of the most common symptoms of StD<sup>27</sup> and, in adolescents, predicts adult MDD<sup>28</sup>. In previous studies, reward dysfunction was observed in depression patients. Pizzagalli et al. found that MDD patients showed significantly weaker responses to gains in the striatum<sup>29</sup>. The abnormal resting-state FCs of the ventral striatum, a key region in the reward network, was observed in StD subjects<sup>30</sup>. Thus, we examined if the identified FCs in the StD classifier were associated with anhedonia and reward responsiveness in the present study.

## Results

**Highly accurate classifier for StD and generalization to an independent cohort.** A classifier was constructed based on the FCs of each subject to distinguish StD subjects from HCs. The classification was first carried out by feature-selection from all 9316 FCs with L1-SCCA and SLR. Then, for this classifier, the weighted linear summation (or linear discriminant function) of the identified FCs was computed. This algorithm was applied to the training dataset ( $n=91$ , Table 1), and then a classification accuracy of 80% (area under the curve 0.84, sensitivity 70%, specificity 85%;  $p=0.001$  with a permutation test, Supplementary Fig. S1a,c) was revealed by the leave-one-out cross validation (LOOCV) procedure. Next, the classifier was tested on an independent validation cohort ( $n=43$ ), and then a classification accuracy of 79% (area under the curve 0.75, sensitivity 84%, specificity 72%;  $p=0.001$  with a permutation test, Supplementary Fig. S1b,d) was revealed.

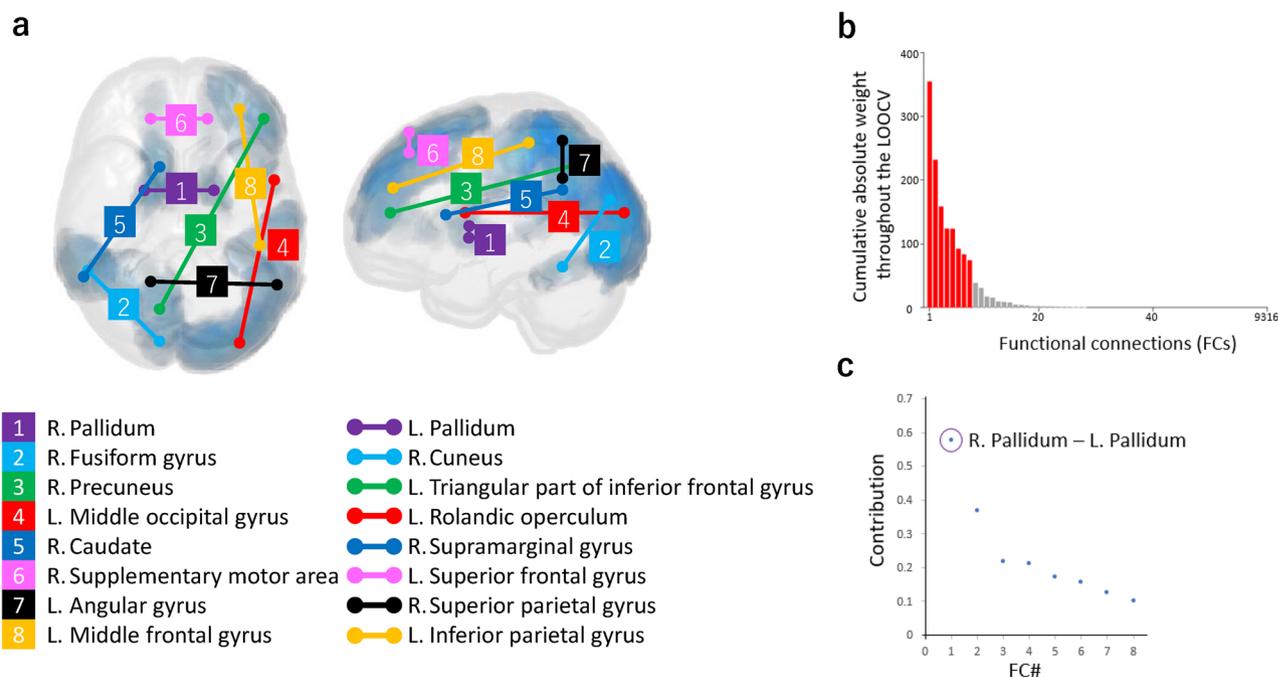
**FCs in the classifier.** Then, the neural basis of the StD classifier was investigated. The sparse classification algorithm allowed the survival of eight FCs from 29 FCs that were selected at least once throughout the LOOCV procedure based on their contribution to the classifier (Table 2 and Fig. 1a). The robustness and stability of the detected FCs throughout the cross-validation procedure are shown in Fig. 1b. Furthermore, we confirmed that the weights of the eight identified FCs across the LOOCV procedure were significantly nonzero (two-sided Wilcoxon signed rank test,  $p<0.001$ ), indicating their important contribution to the classifier. The contribution index of individual FCs to the corresponding connection, which was defined by the difference of each FC between StD subjects and HCs multiplied by the classifier weight, is plotted in Fig. 1c. We found that FC#1 (right pallidum to left pallidum) between inter-hemispheric regions showed the most outstanding contribution of the eight FCs. We also found that only FC#1, out of 9316 FCs, was identified as having a significant difference between StD and HCs by using conventional between-group comparison (two-sample  $t$  test with Bonferroni correction,  $p<0.05/9316$ ).

	Training dataset			Test dataset		
	StD	Healthy controls	$p$ -value	StD	Healthy controls	$p$ -value
No. of participants	30	61	NA	16	27	NA
Sex, male/female	19/11	32/29	0.326	10/6	18/9	0.782
Age (years), mean (SD)	18.2 (0.4)	18.4 (0.5)	0.054	18.6 (0.7)	18.5 (0.6)	0.921
BDI-II, mean (SD)	16.8 (3.5)	3.0 (1.9)	0.000***	20 (5.0)	3 (2.2)	0.000***

**Table 1.** Demographic and clinical information of the participants used to construct the subthreshold depression (StD) classifier. The differences between StD subjects and healthy controls were tested by a two-tailed  $t$  test (for age and BDI-II) or the chi-squared test (for sex). \*\*\* $p<0.001$ . BDI-II Beck's Depression Inventory-II, NA not applicable, SD standard deviation.

ID	Name	Lat	BrainVISA Sulci Atlas (sulcus)	BA	rStD	rHC	Weight	Contribution
1	Pallidum	R	Pallidum	–	0.344	0.513	–3.41	0.578
	Pallidum	L	Pallidum	–				
2	Fusiform gyrus	R	Internal occipito-temporal lateral sulcus	37	0.147	0.341	–1.90	0.370
	Cuneus	R	Cuneal sulcus	18				
3	Precuneus	R	Transverse parietal sulcus	7	–0.200	–0.287	2.51	0.217
	Triangular part of the inferior frontal gyrus	L	Anterior inferior frontal sulcus	45				
4	Middle occipital gyrus	L	Lobe occipital	19	–0.149	–0.038	–1.91	0.213
	Rolandic operculum	L	Anterior sub-central ramus of the lateral fissure	48				
5	Caudate	R	Caudate	–	–0.058	0.044	–1.67	0.172
	Supramarginal gyrus	R	Anterior terminal ascending branch of the superior temporal sulcus	40				
6	Supplementary motor area	R	Median frontal sulcus	6	0.359	0.467	–1.46	0.157
	Superior frontal gyrus	L	Median frontal sulcus	9				
7	Angular gyrus	L	Anterior terminal ascending branch of the superior temporal sulcus	39	–0.153	–0.297	0.88	0.127
	Superior parietal gyrus	R	Superior parietal sulcus	7				
8	Middle frontal gyrus	L	Intermediate frontal sulcus	46	–0.061	0.016	–1.33	0.102
	Inferior parietal gyrus	L	Superior postcentral intraparietal superior sulcus	40				

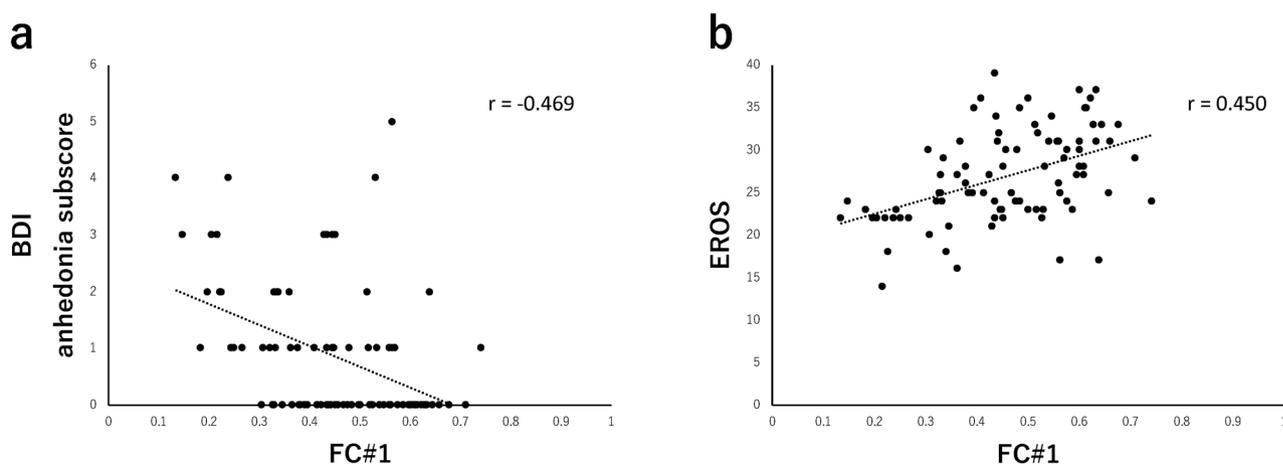
**Table 2.** Characteristics of the eight interregional functional connections (FCs) used in the subthreshold depression (StD) classifier. The mean functional correlation values for StD (rStD) and healthy controls (rHC) are presented for individual FCs. Since the StD group was the positive class in the classification of the StD and HC populations, the sign of weight was positive when  $rStD > rHC$ , whereas it was negative when  $rStD < rHC$ . Contribution was computed as  $(rStD - rHC) \times \text{weight}$ . BA Brodmann area.



**Figure 1.** The eight identified functional connections (FCs) for the subthreshold depression (StD) classifier. (a) The 16 brain regions connected by the eight FCs of the StD biomarker (rendered with: MRICroGL 64-bit, June 12, 2015, <https://www.nitrc.org/projects/mricrogl>). (b) Contribution of individual FCs to the StD biomarker. The cumulative absolute weights are shown for all 9316 FCs, of which 29 FCs were selected at least once throughout the leave-one-out cross validation process. The eight FCs identified through the StD classifier were derived from a key subset of the 29 FCs (the red columns represent the eight identified FCs, and the gray columns represent the remaining 21 FCs). (c) Contribution of each of the eight FCs to the StD biomarker.

Identified FCs	BDI anhedonic subscore		EROS	
	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>
FC#1	-0.469	0.000*	0.450	0.000*
FC#2	-0.321	0.002*	0.282	0.007
FC#3	0.287	0.006	-0.119	0.261
FC#4	-0.264	0.011	0.183	0.083
FC#5	-0.241	0.021	0.185	0.079
FC#6	-0.256	0.014	0.383	0.000*
FC#7	0.199	0.059	-0.217	0.039
FC#8	-0.156	0.141	0.147	0.164

**Table 3.** Relationship between each functional correlation value and clinical score. Clinical scores associated with anhedonia (BDI anhedonic subscore and EROS). \*Bonferroni-adjusted significance level of  $p < 0.05/8$  ( $= 0.006$ ). *BDI* Beck's Depression Inventory, *EROS* Environmental Reward Observation Scale, *FC* functional connection.



**Figure 2.** A scatter plot of correlations between the identified functional connections (FCs) and clinical scores. (a) FC#1 shows a negative correlation with the BDI anhedonic subscore ( $r = -0.469$ ,  $p < 0.006^*$ ). (b) FC#1 shows a positive correlation with EROS score ( $r = 0.450$ ,  $p < 0.006^*$ ). \*Bonferroni-adjusted significance level of  $p < 0.05/8$  ( $= 0.006$ ). *BDI* Beck's Depression Inventory, *EROS* Environmental Reward Observation Scale.

**Association of important FCs with anhedonia and reward responsiveness.** We evaluated the relevance of the important FCs to anhedonia and reward responsiveness, which are features of StD<sup>31</sup>. FC#1 and FC#2 had significant negative correlations with the BDI anhedonic subscore. FC#1 and FC#6 had significant positive correlations with the EROS score. Of the eight FCs, only FC#1 had significant correlations with the BDI anhedonic subscore and EROS score (Table 3, Fig. 2a,b). Also, similar results were obtained in the sum of training and test dataset (Supplementary Table S3).

## Discussion

Herein, we generated a neuroimaging-based biomarker for StD of which the generalization capability was confirmed in an independent validation cohort. Furthermore, our findings suggest that the FC between the left and right pallidum, which was the most important of the eight FCs identified for this classifier, may be associated with clinically important dysfunctions in subjects with StD such as hyposensitivity to rewards and anhedonia.

In recent years, some impaired brain regions and abnormal patterns of brain connectivity have been evaluated in subjects with StD by using rs-fMRI<sup>11–18</sup>. Unfortunately, these findings could not be applied to direct diagnosis, but they are useful for the identification of disease biomarkers associated with the dysfunctions observed in StD. This is because the individual predictive ability of these biomarkers was not evaluated, even though they were statistically significant at the group level. Following this, there has been increasing interest in using machine learning algorithms in order to predict phenotypes in a way that allows characterization at the level of the individual beyond group level analyses<sup>19</sup>. However, also in these studies, there were some limitations regarding the generalizability and feasibility of the identified markers, for example, there were too many predictive FCs to explain StD individually<sup>21</sup>. Yahata et al.<sup>22</sup> overcame these limitations by combining two machine learning algorithms, L1-SCCA and SLR, and identified rs-fMRI-based classifiers for autism spectrum disorder with high generalization<sup>22</sup>. This technique was also used successfully to extract a few FCs related only to the core characteristics of melancholic MDD<sup>25</sup>. By using the same approach, we developed a generalizable rs-fMRI-based

biomarker for StD tested on an independent validation cohort. The test data were acquired using a Siemens Verio scanner that was upgraded to a Skyra scanner including a retrofit of hardware and software. Imaging was also performed with a new protocol that had higher spatial and temporal resolution than used for the discovery data (Supplementary Table S1). Scanner specifications<sup>32</sup> and imaging parameters<sup>33</sup> are known to affect MRI data. Therefore, it is considered to be a great advantage of our biomarker that it showed generalization performance for different machines and different protocols without special harmonization.

Our findings demonstrated that FC#1 (right pallidum to left pallidum), which was the most important of the eight FCs identified for this classifier, may be associated with the clinically important dysfunctions in subjects with StD such as hyposensitivity to rewards and anhedonia. Of the eight identified FCs, FC#1 showed the strongest negative correlation with the BDI anhedonic subscore and strongest positive correlation with the EROS score. In addition, both of these correlations were still significant after multiple comparison corrections only for FC#1. These results suggest that the lower FC between the right and left pallidum may be relevant to hyposensitivity to rewards and anhedonia in subjects with StD. It is difficult to interpret that the FC between bilateral pallidum was not selected as an important feature to discriminate between MDD and healthy subjects in our previous study<sup>25</sup>. To verify whether the key findings of this study targeting StD also apply in MDD, we analyzed previous data (92 patients with MDD vs. 92 HCs from Ichikawa et al.<sup>25</sup>) focusing on the FC between bilateral pallidum. As a result, there was a trend toward decreased FC between bilateral pallidum in MDD compared to HCs, which was not significant ( $p = 0.067$ ). It is unclear why the decreased FC between bilateral pallidum is more characteristic of StD than MDD, but it is possible that the balance of FC changes after onset. This issue needs further clarification in future longitudinal cohort studies.

The right and left pallidum are connected via the anterior commissure through the lateral nucleus of the globus pallidus<sup>34,35</sup>. The ventral pallidum reportedly plays a crucial role in motivation<sup>36</sup>. In rats and monkeys, electrical stimulation within the ventral pallidum supports self-stimulation behavior as a reward in and of itself<sup>37,38</sup>. In addition, in humans, activation in the ventral pallidum is associated with receiving food and money as a reward<sup>39,40</sup>. It has been suggested that the ventral pallidum is deeply involved in the regulation of motor action based on reward expectation<sup>41</sup>. Classically, lesions of the globus pallidus are reportedly related to anhedonia, decreased social interactions, and flattened affect<sup>42,43</sup>. Given that late adolescents with StD have been characterized by reward system dysfunction<sup>26</sup>, our findings suggest that the inter-hemispheric connectivity of the pallidum may be especially important in this process. Indeed, the volume of the bilateral pallidum is reportedly decreased and has a negative association with the severity of depressive symptoms in young StD subjects<sup>44</sup>.

Although the FC between the left and right pallidum was identified as a particularly important biomarker of StD, the other seven FCs contained areas in which abnormalities have been suggested in depression. In particular, the FC between the precuneus and the left triangular part of the inferior frontal gyrus is interesting because it is in the same direction as the connection between the precuneus and left dorsolateral prefrontal cortex, which was identified by Ichikawa et al. as an important connection in the biomarker for melancholic MDD<sup>25</sup>. Further research is needed in longitudinal studies, such as which connections are involved in the development of MDD.

There are four limitations related to the present study. First, because our study was cross-sectional, we cannot determine if these observations are the cause or result of StD. Additionally, the relatively mild depressive symptoms of StD subjects can change easily over time. Future longitudinal studies are required to address this question and will help further our understanding of the neural basis of StD. Second, the sample size of the current study was relatively small. The performance of a classifier is directly influenced by sample size, with which prediction accuracy may decrease<sup>45</sup>. Thus, a larger sample size is needed to validate the current findings further. Third, there were some differences in head motion between the two groups in the training dataset (Supplementary Table S2) that may influence the classifier performance, although six head motion parameters were linearly regressed out. Motion artifacts contribute substantially to the rs-fMRI signal and are not fully countered by motion regressions<sup>58</sup>. Some efforts may have to be directed at preventing head motion, such as behavioral or physical interventions<sup>46</sup>. Moreover, attention should be paid to the neurobiological basis of head motion as identified by Zeng et al.<sup>47</sup>. Differentiating true disease effects from the correlates of motion tendency is critical for using connectivity markers in the clinical area because correlates of motion may reduce specificity of biomarkers. Fourth, the cerebellum was not incorporated in the construction of the classifier because it was truncated in many participants' images. Cerebellar abnormalities have been the focus of increasing concern recently in depression<sup>48</sup>. There are several studies on StD subjects from rs-fMRI that has identified aberrant cerebellar activity<sup>12,13</sup>.

In summary, this study generated a generalizable rs-fMRI-based classifier for the accurate prediction of StD for the first time. For the StD classifier, the FC between the left and right pallidum was the most important of eight FCs identified using sparse classification algorithms, which may be related to clinically important dysfunctions of subjects with StD such as anhedonia and hyposensitivity to rewards. Investigation of whether modulation of the identified FCs can be an effective treatment for StD may be an important future research topic.

## Methods

**Participants.** Participants were enrolled from 18 to 19-year-old first-year students attending Hiroshima University, Japan. Beck's Depression Inventory-II (BDI-II)<sup>49</sup> was administered to the participants to score depressive severity. A BDI-II score of 13 is a suggested cutoff point to detect depression in college students<sup>50,51</sup>; therefore, in the present study, StD was defined as a score of  $\geq 13$  points on the BDI-II. Participants that met the following criteria after completing the Japanese version of the Composite International Diagnostic Interview<sup>52</sup> were excluded from this study: major depressive episode during the past year, a lifetime history of bipolar disorder, taking psychopharmacological or psychological treatment within the past year, possibility of acute suicide attempts, difficulty in understanding the purpose of the study, or difficulty in completing the self-report scales

due to a serious mental condition or severe physical disease. In this way, a total of 91 subjects were enlisted in the rs-fMRI experiments, including 30 StD subjects with BDI-II scores from 13 to 27 (19 men and 11 women) and 61 age-, sex-, and education-matched healthy controls (HCs) with BDI-II scores < 8 (32 men and 29 women). Demographic data for the participants are presented in Table 1. This study was approved by the ethics committee of Hiroshima University, and all participants provided their written informed consent after the study was completely described to each participant. In addition, we confirm that all procedures were carried out in accordance with relevant guidelines and regulations.

**BDI-II.** The BDI-II is a widely-used instrument for measuring depression and depressive symptoms and consists of 21 self-report items that are rated on a 4-point scale ranging from 0 to 3. The Japanese version of the BDI-II has shown good validity and reliability<sup>48</sup>. Additionally, anhedonia was assessed by the BDI-II anhedonic subscore: item #4—loss of pleasure, item #12—loss of interest, item #15—loss of energy, and item #21—loss of sex drive<sup>53–55</sup>.

**Environmental Reward Observation Scale (EROS).** The EROS<sup>56</sup> is used to assess environmental reward and response-contingent positive reinforcement. The original version of this assessment is a 10-item scale measuring agreement (1: strongly disagree to 4: strongly agree). The Japanese version of the EROS has shown good validity and reliability<sup>57</sup>. Additionally, this scale shows a negative correlation with depressive symptoms, especially anhedonia symptoms<sup>57</sup>.

**fMRI data acquisition.** In a darkened scanning room, the participants were required to maintain their gaze on a fixation point in the center of a monitor screen, not to think of anything specific, and to stay awake. Details of the imaging protocols for fMRI data acquisition and the procedure in each dataset are shown in Supplementary Table S1.

**fMRI data preprocessing.** T1-weighted structural images and resting-state functional images were pre-processed with SPM8 (Wellcome Trust Centre for Neuroimaging, University College London, UK) running on MATLAB R2019a (MathWorks, Inc., Natick, MA, USA). The functional images were corrected for slice-timing and subsequently realigned to the mean image. Next, using the associated parameters obtained through the segmentation of the T1-weighted structural images coregistered to the mean functional image, the fMRI data were normalized and resampled in  $2 \times 2 \times 2$  mm<sup>3</sup> voxels. All functional images were then smoothed using an isotropic 6-mm full-width half-maximum Gaussian kernel. Finally, a scrubbing procedure<sup>58</sup> was used to compensate further for motion, removing any volume (i.e., functional images) with excessive movement (frame displacement > 0.5 mm), based on the relative changes from frame to frame in the fMRI time-series (Supplementary Table S2 for a summary of head motion).

**Interregional correlations.** For each participant, pairwise interregional FCs were evaluated between 137 regions of interest, which were defined anatomically using the BrainVISA Sulci Atlas<sup>59,60</sup> covering the entire cerebral cortex and subcortical regions. Then, the time course of fMRI data in each region was extracted. In this study, the cerebellum was not incorporated in the construction of the classifier because it was truncated in many participants' images. After applying a band-pass filter (transmission range, 0.008–0.1 Hz), the temporal fluctuations of the white matter, cerebrospinal fluid, and entire brain as well as six head motion parameters were linearly regressed out. Then, for each participant, we calculated pairwise Pearson's correlations among 137 regions of interest in order to generate a matrix of 9316 FCs, which were normalized using Z-score transformation. The scrubbing procedure was employed to remove any frames exhibiting abrupt head motions in the filtered time course.

**Classification algorithm for FC selection.** An StD classifier with the identified primary FCs was developed through the combined use of two machine learning algorithms. The procedure for selecting characteristic FCs, training a classifier prediction model, and evaluating its generalizability was performed as a sequential process of  $5 \times 5$  nested feature-selection and leave-one-out cross validation (see the schematic flowchart in Supplementary Fig. S2). As the first algorithm, in order to reduce the number of features to remove the effects of nuisance variables (NVs) including sex and age that may result in catastrophic overfitting, L1-SCCA<sup>23</sup> was used. L1-SCCA retained FCs that have an association with a canonical variable related only to the Diagnosis label and not to NVs (details are provided in the Supplementary Information of our previous study<sup>25</sup>). As the second algorithm, in order to identify a small number of features with a high contribution to the classifier, SLR<sup>24</sup> was used. SLR is capable of training a logistic regression model, while objectively deleting FCs that are useless for the purpose of identifying StD. In SLR, the final number of features according to the principle of automatic relevance determination<sup>61</sup> is decided automatically without manually adjusting the hyper-parameters. On this process approach, SLR classifier was trained by using all-but-one subjects in each leave-one-out (LOO) cross-validation (CV). A sequential process of nested-feature selection and LOOCV is applied to prevent information leakage and over-optimistic results<sup>62</sup>. Nested cross validation and LOOCV were combined in our machine learning algorithm. In order to have more than 20 subjects per fold, we used 5-folds CV for StD classifier development. Classification accuracy was evaluated using the output of the logistic regression classifier generated at the end of the LOOCV procedure. In conclusion, through the combined use of two machine learning algorithms, prior to training SLR, it is essential to reduce the input dimension to a certain extent and remove the effects of NVs that may result in catastrophic overfitting. Therefore, before LOOCV, nested feature-selection was carried out using

L1-SCCA. The robustness and stability of the identified FCs across the LOOCV procedure were confirmed by computing their cumulative absolute weights:  $\epsilon^k = \sum_{i=1}^N |\mathcal{W}_i^k|$ , where  $N$  is the number of LOOCV folds (i.e., the number of subjects) and  $\mathcal{W}_i^k$  is the weight associated with the  $k$ -th FC during the  $i$ -th LOOCV fold. A greater magnitude of  $\epsilon^k$  indicates a more significant contribution by the  $k$ -th FC to the classification into StD and HC throughout the LOOCV procedure. More details of this algorithm were provided in a study of autism spectrum disorder<sup>22</sup>. The original code is also accessible (please contact the server administrator of the ATR Brain Information Communication Research Laboratory: asd-classifier@atr.jp).

**Generalization to an independent dataset.** Participants for an independent validation dataset were recruited from first-year students attending Hiroshima University in a different year, who were diagnosed using the BDI-II and Mini-International Neuropsychiatric Interview<sup>63,64</sup>. These participants met the same inclusion and exclusion criteria as the training dataset (StD subjects:  $n = 16$ ; HCs:  $n = 27$ , Table 1). Details of fMRI data acquisition are shown in Supplementary Table S1. The MRI data were preprocessed in the same way as for the training dataset, and then the interregional FCs for each subject were calculated. Written informed consent was obtained from all participants before participating in the study. The ethics committee of Hiroshima University approved this study.

**Correlations between the identified FCs and clinical measures.** The association between the identified FCs in the StD classifier and clinical assessments relative to anhedonia and reward responsiveness (BDI-II anhedonic subscore and EROS, respectively) was tested. For correlation analyses, Pearson's and Spearman's rank correlation coefficients were calculated when normality was not rejected and was rejected, respectively, using SPSS Statistics 25 software (SPSS, Inc., Chicago, IL, USA). Bonferroni's correction was used to adjust for multiple comparisons, and the significance level was set as  $p < 0.05/8$  ( $= 0.006$ ).

### Data availability

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Received: 30 June 2022; Accepted: 6 April 2023

Published online: 18 April 2023

### References

- Cuijpers, P. & Smit, F. Subthreshold depression as a risk indicator for major depressive disorder: A systematic review of prospective studies. *Acta Psychiatr. Scand.* **109**, 325–331 (2004).
- Pincus, H. A., Davis, W. W. & McQueen, L. E. 'Subthreshold' mental disorders. A review and synthesis of studies on minor depression and other 'brand names'. *Br. J. Psychiatry* **174**, 288–296 (1999).
- Mikolajczyk, R. T. *et al.* Prevalence of depressive symptoms in university students from Germany, Denmark, Poland and Bulgaria. *Soc. Psychiatry Psychiatr. Epidemiol.* **43**, 105–112 (2008).
- Kadison, R. College psychiatry 2006: Challenges and opportunities. *J. Am. Coll. Health* **54**, 338–340 (2006).
- Farrer, L. M., Gulliver, A., Bennett, K., Fassnacht, D. B. & Griffiths, K. M. Demographic and psychosocial predictors of major depression and generalised anxiety disorder in Australian university students. *BMC Psychiatry* **16**, 241 (2016).
- Jinnin, R. *et al.* Detailed course of depressive symptoms and risk for developing depression in late adolescents with subthreshold depression: A cohort study. *Neuropsychiatr. Dis. Treat.* **13**, 25–33 (2016).
- Othieno, C. J., Okoth, R. O., Peltzer, K., Pengpid, S. & Malla, L. O. Depression among university students in Kenya: Prevalence and sociodemographic correlates. *J. Affect. Disord.* **165**, 120–125 (2014).
- Ibrahim, A. K., Kelly, S. J., Adams, C. E. & Glazebrook, C. A systematic review of studies of depression prevalence in university students. *J. Psychiatr. Res.* **47**, 391–400 (2013).
- Fergusson, D. M., Horwood, L. J., Ridder, E. M. & Beautrais, A. L. Suicidal behaviour in adolescence and subsequent mental health outcomes in young adulthood. *Psychol. Med.* **35**, 983–993 (2005).
- Klein, D. N., Shankman, S. A., Lewinsohn, P. M. & Seeley, J. R. Subthreshold depressive disorder in adolescents: Predictors of escalation to full-syndrome depressive disorders. *J. Am. Acad. Child. Adolesc. Psychiatry* **48**, 703–710 (2009).
- Philippi, C. L., Motzkin, J. C., Pujara, M. S. & Koenigs, M. Subclinical depression severity is associated with distinct patterns of functional connectivity for subregions of anterior cingulate cortex. *J. Psychiatr. Res.* **71**, 103–111 (2015).
- Wei, X. *et al.* Altered resting-state connectivity in college students with nonclinical depressive symptoms. *PLoS One* **9**, e114603 (2014).
- Wei, X. *et al.* Alteration of spontaneous neuronal activity in young adults with non-clinical depressive symptoms. *Psychiatry Res.* **233**, 36–42 (2015).
- Kaiser, R. H. *et al.* Distracted and down: Neural mechanisms of affective interference in subclinical depression. *Soc. Cogn. Affect. Neurosci.* **10**, 654–663 (2015).
- Petrican, R., Saverino, C., Rosenbaum, R. S. & Grady, C. Inter-individual differences in the experience of negative emotion predict variations in functional brain architecture. *Neuroimage* **123**, 80–88 (2015).
- Gao, C. *et al.* Decreased subcortical and increased cortical degree centrality in a nonclinical college student sample with subclinical depressive symptoms: A resting-state fMRI study. *Front. Hum. Neurosci.* **10**, 617 (2016).
- Ely, B. A. *et al.* Resting-state functional connectivity of the human habenula in healthy individuals: Associations with subclinical depression. *Hum. Brain Mapp.* **37**, 2369–2384 (2016).
- Dedovic, K. *et al.* Psychological, endocrine, and neural correlates of attentional bias in subclinical depression. *Anxiety Stress Coping* **29**, 479–496 (2016).
- Patel, M. J. *et al.* Machine learning approaches for integrating clinical and imaging features in late-life depression classification and response prediction. *Int. J. Geriatr. Psychiatry* **30**, 1056–1067 (2015).
- Orrù, G., Petterson-Yeo, W., Marquand, A. F., Sartori, G. & Mechelli, A. Using Support Vector Machine to identify imaging biomarkers of neurological and psychiatric disease: A critical review. *Neurosci. Biobehav. Rev.* **36**, 1140–1152 (2012).
- Zhu, Y. *et al.* Connectome-based biomarkers predict subclinical depression and identify abnormal brain connections with the lateral habenula and thalamus. *Front. Psychiatry* **10**, 371 (2019).

22. Yahata, N. *et al.* A small number of abnormal brain connections predicts adult autism spectrum disorder. *Nat. Commun.* **7**, 11254 (2016).
23. Witten, D. M., Tibshirani, R. & Hastie, T. A penalized matrix decomposition, with applications to sparse principal components and canonical correlation analysis. *Biostatistics* **10**, 515–534 (2009).
24. Yamashita, O., Sato, M. A., Yoshioka, T., Tong, F. & Kamitani, Y. Sparse estimation automatically selects voxels relevant for the decoding of fMRI activity patterns. *Neuroimage* **42**, 1414–1429 (2008).
25. Ichikawa, N. *et al.* Primary functional brain connections associated with melancholic major depressive disorder and modulation by antidepressants. *Sci. Rep.* **10**, 3542 (2020).
26. Takagaki, K. *et al.* Behavioral characteristics of subthreshold depression. *J. Affect. Disord.* **168**, 472–475 (2014).
27. Rodríguez, M. R., Nuevo, R., Chatterji, S. & Ayuso-Mateos, J. L. Definitions and factors associated with subthreshold depressive conditions: A systematic review. *BMC Psychiatry* **12**, 181 (2012).
28. Pine, D. S., Cohen, E., Cohen, P. & Brook, J. Adolescent depressive symptoms as predictors of adult depression: Moodiness or mood disorder?. *Am. J. Psychiatry* **156**, 133–135 (1999).
29. Pizzagalli, D. A. *et al.* Reduced caudate and nucleus accumbens response to rewards in unmedicated individuals with major depressive disorder. *Am. J. Psychiatry* **166**, 702–710 (2009).
30. Hwang, J. W. *et al.* Enhanced default mode network connectivity with ventral striatum in Subthreshold depression individuals. *J. Psychiatr. Res.* **76**, 111–120 (2016).
31. Takagaki, K. *et al.* Behavioral activation for late adolescents with subthreshold depression: A randomized controlled trial. *Eur. Child Adolesc. Psychiatry* **25**, 1171–1182 (2016).
32. Friedman, L., Glover, G. H., Krenz, D. & Magnotta, V. Reducing inter-scanner variability of activation in a multicenter fMRI study: Role of smoothness equalization. *Neuroimage* **32**, 1656–1668 (2006).
33. Thulborn, K. R., Chang, S. Y., Shen, G. X. & Voyvodic, J. T. High-resolution echo-planar fMRI of human visual cortex at 3.0 tesla. *NMR Biomed.* **10**, 183–190 (1997).
34. Naidich, T. P. *et al.* Anterior commissure: Anatomic-MR correlation and use as a landmark in three orthogonal planes. *Radiology* **158**, 421–429 (1986).
35. Heimer, L., Harlan, R. E., Alheid, G. F., Garcia, M. M. & de Olmos, J. Substantia innominata: A notion which impedes clinical-anatomical correlations in neuropsychiatric disorders. *Neuroscience* **76**, 957–1006 (1997).
36. Smith, K. S., Tindell, A. J., Aldridge, J. W. & Berridge, K. C. Ventral pallidum roles in reward and motivation. *Behav. Brain Res.* **196**, 155–167 (2009).
37. Panagis, G., Miliareisis, E., Anagnostakis, Y. & Spyraiki, C. Ventral pallidum self-stimulation: A moveable electrode mapping study. *Behav. Brain Res.* **68**, 165–172 (1995).
38. Bichot, N. P., Heard, M. T. & Desimone, R. Stimulation of the nucleus accumbens as behavioral reward in awake behaving monkeys. *J. Neurosci. Methods* **199**, 265–272 (2011).
39. Beaver, J. D. *et al.* Individual differences in reward drive predict neural responses to images of food. *J. Neurosci.* **26**, 5160–5166 (2006).
40. Pessiglione, M. *et al.* How the brain translates money into force: A neuroimaging study of subliminal motivation. *Science* **316**, 904–906 (2007).
41. Tachibana, Y. & Hikosaka, O. The primate ventral pallidum encodes expected reward value and regulates motor action. *Neuron* **76**, 826–837 (2012).
42. Bhatia, K. P. & Marsden, C. D. The behavioural and motor consequences of focal lesions of the basal ganglia in man. *Brain* **117**, 859–876 (1994).
43. Laplane, D. *et al.* Obsessive-compulsive and other behavioural changes with bilateral basal ganglia lesions. A neuropsychological, magnetic resonance imaging and positron tomography study. *Brain* **112**, 699–725 (1989).
44. Li, J. *et al.* Anatomical brain difference of subthreshold depression in young and middle-aged individuals. *Neuroimage Clin.* **14**, 546–551 (2017).
45. Arbabshirani, M. R., Plis, S., Sui, J. & Calhoun, V. D. Single subject prediction of brain disorders in neuroimaging: Promises and pitfalls. *Neuroimage* **145**, 137–165 (2017).
46. Power, J. D. A simple but useful way to assess fMRI scan qualities. *Neuroimage* **154**, 150–158 (2017).
47. Zeng, L. L. *et al.* Neurobiological basis of head motion in brain imaging. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 6058–6062 (2014).
48. Wang, L., Hermens, D. F., Hickie, I. B. & Lagopoulos, J. A systematic review of resting-state functional-MRI studies in major depression. *J. Affect. Disord.* **142**, 6–12 (2012).
49. Kojima, M. & Furukawa, T. *Japanese Version of the Beck Depression Inventory* 2nd edn. (Nippon-Hyoron-sha Co., 2003).
50. Dozois, D. J. A., Dobson, K. S. & Ahnberg, J. L. A psychometric evaluation of the Beck Depression Inventory-II. *Psychol. Assess.* **10**, 83–89 (1998).
51. Wang, Y. P. & Gorenstein, C. Psychometric properties of the Beck Depression Inventory-II: A comprehensive review. *Braz. J. Psychiatry* **35**, 416–431 (2013).
52. Kawakami, N. *et al.* Twelve-month prevalence, severity, and treatment of common mental disorders in communities in Japan: Preliminary finding from the World Mental Health Japan Survey 2002–2003. *Psychiatry Clin. Neurosci.* **59**, 441–452 (2005).
53. Pizzagalli, D. A. *et al.* Toward an objective characterization of an anhedonic phenotype: A signal-detection approach. *Biol. Psychiatry* **57**, 319–327 (2005).
54. Joiner, T. E., Brown, J. S. & Metalsky, G. I. A test of the tripartite model's prediction of anhedonia's specificity to depression: Patients with major depression versus patients with schizophrenia. *Psychiatry Res.* **119**, 243–250 (2003).
55. Treadway, M. T., Buckholtz, J. W., Schwartzman, A. N., Lambert, W. E. & Zald, D. H. Worth the “Effort”? The effort expenditure for rewards task as an objective measure of motivation and anhedonia. *PLoS One* **4**, e6598 (2009).
56. Armento, M. E. & Hopko, D. R. The Environmental Reward Observation Scale (EROS): Development, validity, and reliability. *Behav. Ther.* **38**, 107–119 (2007).
57. Kunisato, Y. *et al.* Development of Japanese version of Environmental Reward Observation Scale (EROS). *Jpn. J. Behav. Ther.* **37**, 21–31 (2011).
58. Power, J. D., Barnes, K. A., Snyder, A. Z., Schlaggar, B. L. & Petersen, S. E. Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *Neuroimage* **59**, 2142–2154 (2012).
59. Perrot, M., Rivière, D. & Mangin, J. F. Cortical sulci recognition and spatial normalization. *Med. Image Anal.* **15**, 529–550 (2011).
60. Rivière, D. *et al.* Automatic recognition of cortical sulci of the human brain using a congregation of neural networks. *Med. Image Anal.* **6**, 77–92 (2002).
61. Tipping, M. E. & Faul, A. Fast marginal likelihood maximisation for sparse Bayesian models. *Proceedings of the Ninth International Workshop on Artificial Intelligence and Statistics* (2003).
62. Whelan, R. & Garavan, H. When optimism hurts: Inflated predictions in psychiatric neuroimaging. *Biol. Psychiatry* **75**, 746–748 (2014).
63. Sheehan, D. V. *et al.* The Mini-International Neuropsychiatric Interview (M.I.N.I.): The development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *J. Clin. Psychiatry* **59**, 22–33 (1998).
64. Otsubo, T. *et al.* Reliability and validity of Japanese version of the Mini-International Neuropsychiatric Interview. *Psychiatry Clin. Neurosci.* **59**, 517–526 (2005).

## Acknowledgements

This research was supported by AMED under Grant Numbers JP18dm0307002, JP18dm0307008; JST-Mirai Program under Grant Number JPMJMI20D6; and JSPS Grant-in-Aid for Scientific Research (B) under Grant Number 20H03603. The funders had no role in study design, data collection, data analysis, data interpretation, or writing of the manuscript. A draft of this manuscript was edited by NAI, Inc. (Yokohama, Japan).

## Author contributions

Y.S., G.O., S.Y., N.I. and Y.O. designed the study, collected the data, performed analysis, drafted the manuscript, discussed the results and conclusions for editing the manuscript. All authors contributed to the article and approved the submitted version.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-023-33077-3>.

**Correspondence** and requests for materials should be addressed to Y.O.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023