Research Paper

Classifying youth with bipolar disorder versus healthy youth using cerebral blood flow patterns

Nicholas J. Luciw, PhD; Anahit Grigorian, MSc; Mikaela K. Dimick, PhD; Guocheng Jiang, BSc; J. Jean Chen, PhD; Simon J. Graham, PhD; Benjamin I. Goldstein, MD, PhD; Bradley J. MacIntosh, PhD

Background: Clinical neuroimaging studies often investigate group differences between patients and controls, yet multivariate imaging features may enable individual-level classification. This study aims to classify youth with bipolar disorder (BD) versus healthy youth using grey matter cerebral blood flow (CBF) data analyzed with logistic regressions. **Methods:** Using a 3 Tesla magnetic resonance imaging (MRI) system, we collected pseudo-continuous, arterial spin-labelling, resting-state functional MRI (rfMRI) and T_1-weighted images from youth with BD and healthy controls. We used 3 logistic regression models to classify youth with BD versus controls, controlling for age and sex, using mean grey matter CBF as a single explanatory variable, quantitative CBF features based on principal component analysis (PCA) or relative (intensity-normalized) CBF features based on PCA. We also carried out a comparison analysis using rfMRI data. **Results:** The study included 46 patients with BD (mean age 17 yr, standard deviation [SD] 1 yr; 25 females) and 49 healthy controls (mean age 16 yr, SD 2 yr; 24 females). Global mean CBF and multivariate quantitative CBF offered similar classification performance that was above chance. The association between CBF images and the feature map was not significantly different between groups (p = 0.13); however, the multivariate classifier identified regions with lower CBF among patients with BD (ΔCBF = –2.94 mL/100 g/min; permutation test p = 0047). Classification performance decreased when considering rfMRI data. **Limitations:** We cannot comment on which CBF principal component is most relevant to the classification. Participants may have had various mood states, comorbidities, demographics and medication records. **Conclusion:** Brain CBF features can classify youth with BD versus healthy controls with above-chance accuracy using logistic regression. A global CBF feature may offer similar classification performance to distinct multivariate CBF features.

Introduction

Bipolar disorder (BD) is a severe psychiatric condition defined by recurrent episodes of elevated and depressed mood. The onset of BD most commonly occurs during adolescence.[1] Diagnosis of BD relies on clinical interview and observation, and is often determined by recall of previous illness episodes. Diagnostic reliability for BD is good in adults and youth, but opportunities for improvement exist, especially for youth.[2] Distinguishing between severe mental illnesses among youth is difficult. Although nonoverlapping symptoms can be helpful in this distinction, these illnesses have many overlapping symptoms, and adding objective imaging measures to standard clinical measures has the potential to improve classification, compared with using clinical measures alone.[3] Early intervention is important for reducing symptomatic burden, improving functioning and mitigating illness progression.[4] As such, strategies for early identification are needed to enable early interventions and, in turn, optimize treatment efficacy.[5] Additional research integrating clinical measures with neuroimaging measures and machine learning methods is needed to advance toward the goals of accurate classification and early identification.

Using magnetic resonance imaging (MRI), we may better understand group-level differences in brain anatomy and function among adults and youth with and without BD.[6–8] Such group-level findings have prompted predictive classification and machine learning analyses that aimed to uncover phenotypes of BD in individual neuroimages rather than group-level aggregates, with the ultimate goal of developing diagnostic or prognostic measures. The feasibility of imaging-based classification of adults with BD from healthy controls has been shown using neuroanatomical T_1-weighted MRI; accuracies of 45%–81% have been reported for single-site studies[9–11] and 65% accuracy was reported in a large multisite study.[11] Previously, cerebral blood flow (CBF) estimates from the subgenual cingulate gyrus led to 81% accuracy in classifying bipolar depression versus unipolar depression.[12] The

Correspondence to: B.J. MacIntosh, 2075 Bayview Avenue, Room M6-180, Toronto, Ont., M4N 3M5; brad.macintosh@utoronto.ca

Submitted Jan. 13, 2023; Revised Apr. 14, 2023; Revised May 26, 2023; Accepted May 27, 2023

Cite as: *J Psychiatry Neurosci* 2023 August 29;48(4). doi: 10.1503/jpn.230012

success of imaging-based classification depends on various factors, such as the image quality and sources of image contrast. Magnetic resonance imaging can yield multiple forms of imaging contrast, including neuroanatomical information, as well as neurovascular or hemodynamic information, through arterial spin labelling (ASL) or blood oxygenation level–dependent (BOLD) imaging.

We sought to use ASL-based CBF maps to classify youth with and without BD. Cerebral blood flow is a fundamental aspect of physiology and reflects disease-related changes in brain function and physiology; therefore, CBF may supplement anatomic MRI readouts in classification schemes. A neurovascular measure, namely CBF, is particularly relevant in BD, which is thought to be related to the vascular diseases.[13] Previous studies have found that youth with BD had higher average CBF in grey matter and higher regional CBF in the cingulate gyrus and frontal regions, compared with healthy controls.[14–16] Cerebral blood flow may provide classification capacity to neuroanatomical imaging. Whether global average CBF or multivariate regional CBF is the better choice for classifying individual youth with BD remains questionable. We sought to compare univariate and multivariate CBF as explanatory variables in separate classification models based on logistic regression and features derived from principal component analysis (PCA).[17,18] With a paucity of CBF-based BD classification studies, we chose this parsimonious approach as a proof-of-concept to avoid possible spurious associations (i.e., overfitting). We hypothesized that multivariate CBF features would enable better-than-chance classification of youth with BD and healthy controls.

Methods

Participants

We recruited study participants through community advertising and set up a specialty clinic at Sunnybrook Health Sciences Centre in Toronto, Canada. The exclusion criteria included inability to provide informed consent; cardiac, autoimmune or inflammatory illness; neurologic or cognitive impairment; current anti-inflammatory, antiplatelet, antilipidemic, antihypertensive or hypoglycemic agents; and contraindication to MRI (e.g., ferromagnetic implants, claustrophobia). We established psychiatric diagnoses using the Schedule for Affective Disorders and Schizophrenia for School-Age Children, Present and Lifetime Version (K-SADS-PL).[19] In the standard mood sections, we selected the K-SADS Depression Rating Scale and the K-SADS Mania Rating Scale.[20,21] Participants with BD could meet the criteria for BD-I, BD-II or BD-NOS (not otherwise specified) subtypes, where BD-NOS was defined using operationalized criteria from the Course and Outcome of Bipolar Illness in Youth study.[22] For each participant, we also conducted a Family History Screen interview and a Children's Global Assessment Scale.[23,24] Healthy control participants had no lifetime history of mood or psychotic disorders and no anxiety disorders or alcohol or drug dependence within 3 months of recruitment.

We measured participants' height and weight in light clothing to the nearest 0.5 cm and 0.1 kg following the anthropometry procedures of the National Health and Nutrition Examination Survey.[25] We adjusted weight for clothing choice (1.4 kg for long pants and long-sleeved shirt, 1.1 kg for short pants or short-sleeved shirt, 0.9 kg for short pants and short-sleeved shirt). We also recorded the blood pressure using an automated monitor while seated after 10 minutes of rest (Omron).

MRI acquisition

We used a 3 Tesla MRI system (Achieva, Philips) to generate brain images. An 8-channel head coil receiver and paddings were equipped for patient comfort and to limit head motion. We acquired anatomic T_1-weighted, ASL and BOLD images. The T_1-weighted image used a fast field echo acquisition (repetition time 9.5 ms, echo time 2.3 ms, spatial resolution 0.9 × 0.7 × 1.2 mm, field of view 240 × 191 × 168 mm, scan duration 8 min and 56 s). We performed the ASL scan with a pseudo-continuous labelling scheme (labelling duration 1650 ms, post-label delay for the most inferior slice 1600 ms, 35 control–label pairs) and a single-shot 2-dimensional echo planar imaging acquisition (repetition time 4000 ms, echo time 9.6 ms, spatial resolution 3 × 3 × 5 mm, field of view 192 × 192 × 90 mm, scan duration 4 min and 48 s). We also acquired the proton-density images for ASL calibration (repetition time 10 s, acquisition otherwise identical to ASL). Finally, we collected BOLD fMRI data with T_2*-weighted contrast using echo planar imaging (repetition time 1500, echo time 30 ms, spatial resolution 3 × 3 × 4 mm, field of view 230 × 181 mm, and volumes 231 for a scan duration of 5 min and 46 s).

MRI processing

We calculated the CBF maps from ASL images using a processing pipeline built with tools from FSL (FMRIB Software Library).[26] Pipeline steps included motion correction, control–label image subtraction, smoothing and registration to standard image space. We calculated difference images using sinc-interpolated subtraction of control–label images, spatially smoothed using an isotropic Gaussian kernel with a full width at half maximum of 5 mm.[27] We discarded images affected by excessive head motion using the ENABLE (Enhancement of Automated Blood Flow Estimates) quality control algorithm.[28] We converted the CBF-weighted images into mL/100 g/min units by scaling with the proton density image and calibrating with an established ASL MRI model, using literature values for model parameters.[29] To facilitate a sensitivity analysis of the CBF maps, we corrected the partial volume effect using a regression algorithm (obtained from the ExploreASL software package) that treated the intensity of each voxel as the sum of the signal contribution from each tissue type (i.e., grey and white matter).[30,31]

We used the T_1-weighted images to register the CBF images with the MNI152 template. We then segmented T_1-weighted images to obtain a grey matter segmentation using the FSL-FAST tool.[32] We removed voxels adjacent to lateral

ventricles using an overlay of the right and left lateral ventricle based on the Harvard–Oxford atlas (distributed with FSL, https://fsl.fmrib.ox.ac.uk/). We excluded masked voxels if more than 2 participants did not have ASL coverage in the voxel. For the BOLD fMRI data, we created fractional amplitude of low-frequency oscillation (fALFF) maps using an established pipeline.[33] For each voxel in fMRI images, we computed the square root of the power spectrum to calculate the amplitude value for each voxel. We calculated fALFF by summing the amplitude data in each voxel, which falls in the 0.01–0.08 Hz low-frequency range, and dividing by the sum of amplitude in the entire frequency spectrum.[33] We accounted for the effects of age and sex by regressing these variables against voxelwise maps across all participants in the study to produce adjusted images for all subsequent CBF and fALFF analyses.

Imaging-based classification using logistic regression

We built classification models based on binary logistic regression and fit them to the imaging data using group (healthy control or BD) as the dependent variable and adjusted CBF (or fALFF) as the explanatory variable. We used principal component analysis (PCA) to extract multivariate features representing regions with positively and negatively covarying CBF across the group; these principal components are uncorrelated and are suitable as consecutive variables in the regression. Importantly, principal components are calculated agnostic to the participant diagnostic group.

We built 3 logistic regression models to classify youth with BD versus controls, controlling for age and sex. Model 1 used mean grey matter CBF as a single explanatory variable. Model 2 used quantitative CBF features based on PCA. Model 3 used relative (intensity-normalized) CBF or fALFF features based on PCA. The first approach involved a single explanatory variable, whereas the second and the third approaches involved multiple explanatory variables. For the multivariate logistic regressions, we used stepwise logistic regression to select the consecutive principal components that best explained group membership. The principal components were calculated using a participant-by-voxel data matrix, ordered by the percentage of explained variance. We then performed forward stepwise regression to limit the number of explanatory features and avoid overfitting. The stepwise regression started solely with the first principal component. We added successive principal components until 15 were considered, corresponding to the total number of explanatory variables in previous models.[18,34] The most suitable model was defined a priori by the lowest value of the Akaike information criterion (AIC).[35]

We ran the logistic regression models with repeated 6-fold cross-validation, with 10 repeats, to reduce variance in the model performance estimate. We chose 6 folds to approximate the conventional 80:20 train-to-test split in each fold, and we assigned a roughly even balance of BD and healthy controls to each test set. We assessed the classification performance of each test set using the accuracy value (defined as the percent of participants correctly classified) and the area under the receiver operating characteristic curve (area under the curve [AUC]) for each cross-validation fold. Of the 2 multivariate feature sets, we selected the set with the highest combination of AUC and accuracy for further analysis.

Consolidation of components into a composite bipolar disorder–related pattern

The coefficients from the logistic regression of CBF principal components represent the relevance of each principal component in classifying the 2 groups. The linear combination of the principal components with these coefficients as weights returns a composite CBF pattern related to CBF features that distinguish youth with BD from healthy controls. For models 2 and 3, we quantified the expression of the composite pattern by calculating the dot product of the *z*-scored composite pattern with the CBF image of the participant (i.e., the sum of all voxels after multiplying the 2 images). We then assessed the stability of the voxel estimates in the composite pattern by repeating the entire composite pattern calculation (including PCA and stepwise logistic regression) 1000 times using a bootstrapped resampling of the participants. A *z*-score image indicated the mean and standard deviation of the 1000 patterns. We assessed the significance of the voxels in this image using cluster-based inference via FSL's cluster tool, with a primary *z*-score threshold of 3.29 and a secondary *p*-value threshold of 0.01, as suggested by Woo and colleagues.[36]

Region-of-interest analysis for the CBF multivariate models

To provide insight into the CBF group differences associated with the significant voxels in the composite CBF pattern, we performed a region-of-interest analysis using the composite pattern derived from the best-performing logistic regression model. We considered 3 sets of voxels in the group-difference composite pattern, namely those that were positive (i.e., BD > healthy controls), negative (i.e., BD < healthy controls) and the remaining nonsignificant grey matter voxels. For each of these masks, we calculated the difference between the mean relative CBF (i.e., CBF value normalized to a mean of 80 mL/100 g/min in each image) of the BD and healthy control groups. We chose the relative CBF for the region-of-interest analysis because the PCA analysis was designed to assess between-region relationships rather than a global CBF effect.[37]

Statistical analysis

We conducted a power analysis to compute an anticipated effect size and power, given our current sample size and assuming a univariate analysis of CBF data. We used G*power software and relied on a 2-tailed *t* test to infer CBF differences between 2 independent groups. Based on assumptions of an α value of 0.05, power of 0.80, total sample size of 95 and a univariate statistic, we deduced that we would need an effect size of 0.51 to produce a statistically significant group difference. For group differences in mean grey matter CBF, we used a 2-tailed *t* test, controlling for age and sex, to assess the statistical significance (a priori $\alpha = 0.05$). For group differences in

Table 1 (part 1 of 2): Demographic and clinical characteristics of participants

Characteristic	No. (%) of participants*		t or χ^2	p value
	Healthy controls n = 49	Bipolar disorder n = 46		
Demographics				
Age, yr, mean ± SD	16.30 ± 1.75	17.30 ± 1.23	–3.24	0.002
Sex, female	24 (49.0)	25 (54.3)	0.10	0.75
Non-White	22 (44.9)	14 (30.4)	1.88	0.17
BMI, mean ± SD	20.98 ± 2.81	23.81 ± 3.71	–4.21	< 0.001
Intact family	32 (65.3)	28 (60.9)	0.055	0.81
Socioeconomic status, HI score, mean ± SD	4.35 ± 0.86	4.17 ± 0.90	0.56	0.45
Tanner stage, mean ± SD	4.16 ± 0.59	4.33 ± 0.67	3.00	0.08
CGAS: Most severe past episode, mean ± SD	79.0 ± 7.9	41.9 ± 8.3	7.52	< 0.001
CGAS: Highest past year, mean ± SD	90.4 ± 5.1	67.5 ± 11.7	12.1	< 0.001
CGAS: Past month, mean ± SD	90.7 ± 4.3	64.0 ± 12.1	14.0	< 0.001
Lifetime clinical characteristics				
BD-I	0	15 (32.6)	—	—
BD-II	0	15 (32.6)	—	—
BD-NOS	0	16 (34.8)	—	—
Age of BD onset, yr, mean ± SD	0	14.7 ± 2.15	—	—
Psychosis	0	16 (34.8)	—	—
Suicide attempts	0	7 (15.2)	—	—
Self-injurious behaviour	0	23 (50.0)	—	—
Suicidal ideation	3 (12.5)	29 (63.0)	—	—
Physical or sexual abuse	1 (2.0)	2 (7.2)	—	—
Psychiatric hospital admission	0	25 (54.3)	—	—
Current depression score, mean ± SD	0.41 ± 1.47	14.15 ± 11.18	—	—
Lifetime depression score, mean ± SD	1.20 ± 2.56	29.93 ± 12.56	—	—
Current mania score, mean ± SD	0.08 ± 0.34	9.26 ± 10.12	—	—
Lifetime mania score, mean ± SD	0.63 ± 1.83	30.57 ± 10.97	—	—
Lifetime comorbid diagnoses				
Attention-deficit/hyperactivity disorder	6 (12.2)	22 (47.8)	—	—
Oppositional defiant disorder	0	11 (23.9)	—	—
Conduct disorder	0	2 (4.3)	—	—
Any anxiety disorder	1 (2.0)	35 (76.1)	—	—
No. of anxiety disorders, mean ± SD	0.04 ± 0.29	1.57 ± 1.42	—	—
Substance use disorder	0	12 (26.1)	—	—
Eating disorder	0	10 (21.7)	—	—
Nicotine use	3 (6.1)	20 (43.5)	—	—
Family psychiatric history				
Mania or hypomania	0	24 (55.8)	—	—
Depression	8 (16.3)	31 (72.1)	—	—
Suicide attempt	5 (10.2)	16 (37.2)	—	—
Anxiety	7 (14.3)	27 (62.8)	—	—
Psychosis	0	11 (25.6)	—	—
Substance use disorder	4 (8.2)	21 (45.7)	—	—
Attention-deficit/hyperactivity disorder	3 (6.1)	10 (23.3)	—	—

composite CBF patterns, we compared the scores representing the similarity between the composite CBF pattern and each participant's CBF image between healthy controls and participants with BD using a permutation test. We repeated the entire composite pattern calculation 1000 times (including PCA and stepwise logistic regression), each with a different random permutation of the group labels. For each new composite pattern, we calculated the difference between the mean score of the 2 groups; these differences formed the null distribution used to assign significance of the observed difference in a single-tailed fashion (a priori $\alpha = 0.05$), as we designed the analysis to test whether the group with BD had a higher score. We also conducted a within-group linear model to test for an association between mean grey matter CBF and current

Table 1 (part 2 of 2): Demographic and clinical characteristics of participants

Characteristic	No. (%) of participants*		t or χ^2	p value
	Healthy controls n = 49	Bipolar disorder n = 46		
Lifetime medications				
Second-generation antipsychotics	0	34 (73.9)	—	—
Lithium	0	12 (26.1)	—	—
SSRI antidepressants	0	16 (34.8)	—	—
Non-SSRI antidepressants	0	7 (15.2)	—	—
Stimulants	4 (8.2)	9 (19.6)	—	—
Lamotrigine	0	10 (21.7)	—	—
Valproate	0	2 (4.3)	—	—
Current medications				
Second-generation antipsychotics	0	30 (65.2)	—	—
Lithium	0	10 (21.7)	—	—
SSRI antidepressants	0	4 (8.7)	—	—
Non-SSRI antidepressants	0	3 (6.5)	—	—
Stimulants	3 (6.1)	2 (4.3)	—	—
Lamotrigine	0	10 (21.7)	—	—
Valproate	0	0	—	—
Any current medication	3 (6.1)	39 (84.8)	—	—

BD-I = bipolar disorder subtype I; BD-II = bipolar disorder subtype II; BD-NOS = bipolar disorder not otherwise specified; BMI = body mass index; CGAS = Children's Global Assessment Scale; HI = Hollingshead Index; SD = standard deviation; SSRI = selective serotonin reuptake inhibitor.

depression or mania score. Finally, we estimated the confidence intervals (CIs) of the mean group differences in the region-of-interest analysis of relative CBF using bootstrapping. We resampled CBF images with replacement, and recalculated the mean regional CBF difference between the groups 5000 times. We estimated the 95% CI as the 2.5th and 97.5th percentiles of this bootstrapped distribution, and the p values for these region-of-interest analyses using a permutation test (5000 shuffles of group membership labels).

Ethics approval

The Research Ethics Board at Sunnybrook Research Institute in Toronto, Canada, approved the study (no. 408-2011 and 409-2013). All participants provided free and informed consent to participate in the study.

Results

The study included 95 youth with (n = 46) or without (n = 49) BD. Table 1 lists the demographic and clinical information of the study participants. Participants with BD were slightly older and had a higher body mass index than healthy controls. Most participants with BD were taking medication at the time of the study, most commonly second-generation antipsychotics. Figure 1A shows the mean CBF map of the healthy control group and is also representative of the mean CBF map of the BD group. There was a significant group effect on mean grey matter CBF after controlling for age and sex (t = 2.4, p = 0.018). The mean grey matter CBF was 77.8 (standard deviation [SD] 13.0) mL/100 g/min in the healthy control group, and 83.2 (SD 14.9) mL/100 g/min in the BD group.

Imaging-based regressions, principal components and classification

From the within-group linear model for the BD group, we found that mean grey matter CBF was significantly associated with the current depression score (p = 0.038) but not the current mania score (p = 0.702). For the control group, mean grey matter CBF was not associated with either of these symptomatology scores (p = 0.911 and 0.162 for depression and mania scores, respectively). The results are available in Appendix 1, Table 1, available at www.jpn.ca/lookup/doi/10.1503/jpn.230012/tab-related-content.

Table 2 shows the performance metrics averaged across all folds of repeated cross-validation for the logistic regression models with the lowest AIC. Models 1 and 2 achieved classification accuracy values that exceeded chance (i.e., 95% CIs did not overlap with 0.5), whereas model 3, which used relative CBF multivariate features, had lower performance. For comparison, using adjusted fALFF maps did not yield a classification accuracy greater than chance (mean accuracy 0.56, SD 0.12; Appendix 1, Table 2). The classification accuracy for the mean grey matter CBF (mean accuracy 0.60, SD 0.11) and CBF principal components were similar (mean accuracy 0.61, SD 0.10). The classification accuracy for the relative CBF principal components (mean accuracy 0.52, SD 0.10) was lower than the model using quantitative CBF principal components. The percent variance explained by the principal components calculated from quantitative CBF is shown in Figure 1B. Cumulatively, the principal components explained 57.0% of the variance in CBF across participants. The AIC from each model in the stepwise regression is shown in Figure 1C.

Composite bipolar disorder–related CBF pattern

Figure 2A shows the composite CBF pattern calculated from quantitative CBF (model 2). Positive loadings were found in the right putamen, bilateral nucleus accumbens, right insula, left hippocampus, right caudate, right frontal pole and right superior occipital gyrus. Negative loadings were found in the precuneus, anterior cingulate gyrus, posterior cingulate gyrus, left precentral gyrus and left superior–middle frontal gyrus. We did not find a statistically significant difference in the expression of this pattern between participants with BD and healthy controls ($t = 1.53$, $p = 0.13$; Figure 2B).

Region-of-interest analysis based on composite CBF pattern

Figures 2C and 2D shows group CBF differences for the regions with positive loadings in the composite pattern and those with negative loading. The CBF was lower among participants with BD for regions with positive loadings (ΔCBF = 2.4 mL/100 g/min, 95% CI 0.3 to 5.8 mL/100 g/min; permutation test $p = 0.047$). No significant group difference was observed in regions with negative loadings (ΔCBF = –2.9 mL/100 g/min, 95% CI –5.9 to –0.2 mL/100 g/min; permutation test $p = 0.060$). No significant difference was observed in the remaining regions (ΔCBF = 0.1 mL/100 g/min, 95% CI –0.1 to 0.3 mL/100 g/min; permutation test $p = 0.20$).

Discussion

In this study, we showed the feasibility of classifying youth with and without BD using logistic regression models that relied on univariate and multivariate quantitative CBF imaging features. The logistic regression models that retained quantitative CBF information provided meaningful classification accuracy, whereas the relative CBF and fALFF approaches were inferior and did not classify better than chance. We consolidated multivariate CBF patterns into a single imaging phenotype related to CBF group differences. The loadings of this composite CBF pattern enabled region-of-interest analysis, which showed an interesting pattern of decreased striatum CBF in the BD group (i.e., in the caudate, putamen and nucleus accumbens regions).

This study showed that both univariate and multivariate analyses of CBF data can be used for better-than-chance classification of youth with BD. We observed higher performance than was observed by Almeida and colleagues[12] using support vector machine analysis of CBF in adult

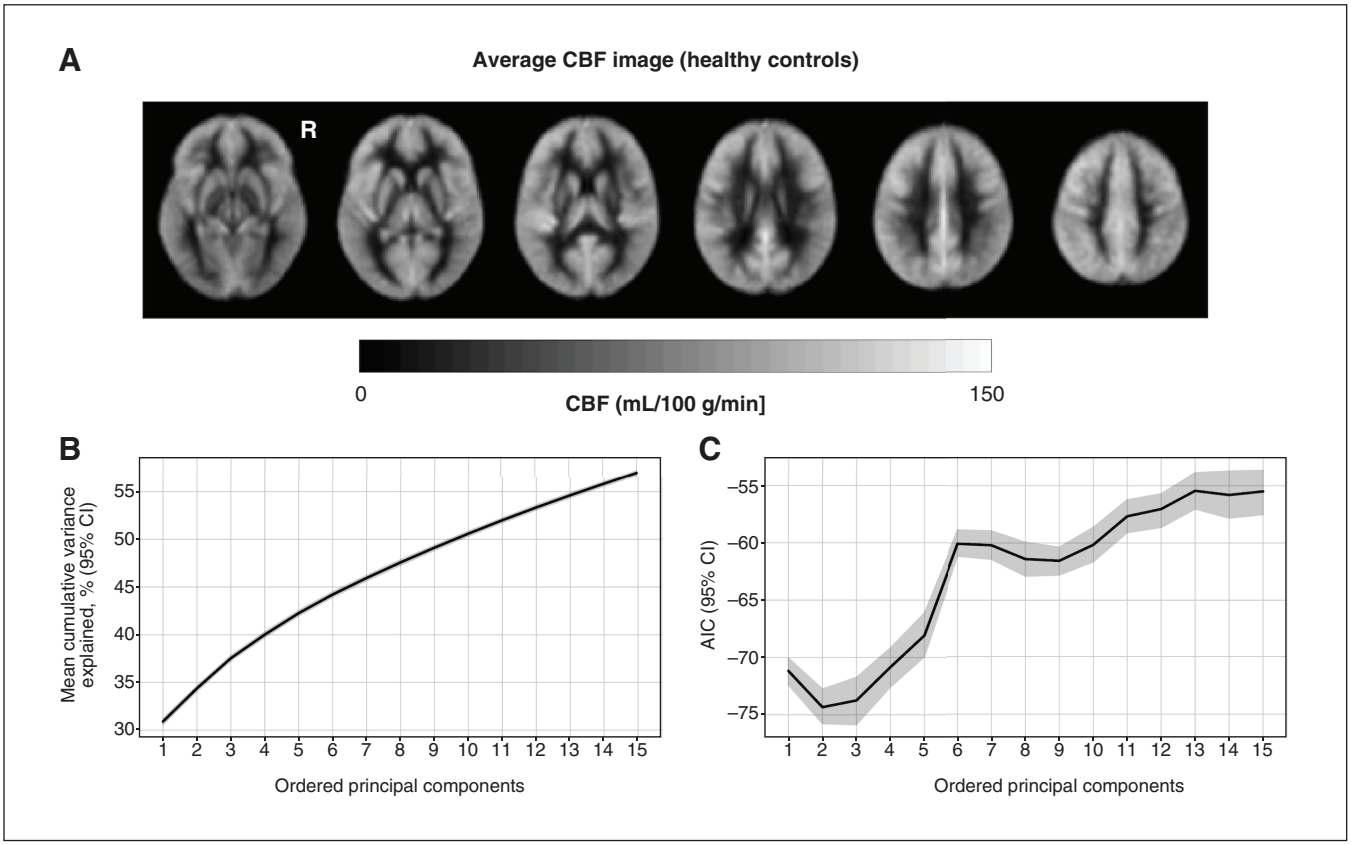

Figure 1: (A) Group average cerebral blood flow (CBF) image for the healthy controls; this was visually representative of the group with bipolar disorder. (B) Percent variance explained by the principal components, calculated from quantitative CBF data. Solid line shows mean across cross validation folds and translucent band shows 95% confidence intervals (CIs); note that the CI closely overlaps with the mean. (C) Akaike information criteria (AIC), as a function of the number of included principal components, is shown for logistic regression models fit using multivariate information from the quantitative CBF data. Lower AIC values indicate a better model fit. R = right hemisphere.

females with BD. Performance was comparable to classification analyses using structural MRI in adult BD, with 45%–81% accuracy in single-site studies.[9,11] The high signal-to-noise ratio of univariate analysis of mean grey matter CBF is desirable for repeatability, whereas multivariate methods give insight into the regional neurophysiological correlates of youth BD. Using relative CBF (i.e., normalized to the same grey matter means in every image) produced poor classification, suggesting that the global grey matter CBF is highly pertinent for classification of BD. Such inter-individual variability is key in the study of CBF in youth BD. Global mean and regional CBF features contributed to our ability to classify adolescents with or without BD. It is important to place this finding in the context of potential mechanistic underpinnings, and several factors can contribute to CBF differences. On the one hand, bioenergetic

Table 2: Performance of logistic regression classification of diagnostic group

	Mean ± SD	
CBF feature	AUC	Accuracy
Mean grey matter CBF	0.62 ± 0.13*	0.60 ± 0.11*
Quantitative CBF principal components	0.60 ± 0.12*	0.61 ± 0.10*
Relative CBF principal components	0.56 ± 0.12*	0.52 ± 0.10

AUC = area under the receiver operating characteristic curve; CBF = cerebral blood flow; SD = standard deviation.
*Indicates the 95% confidence interval does not contain 0.50.

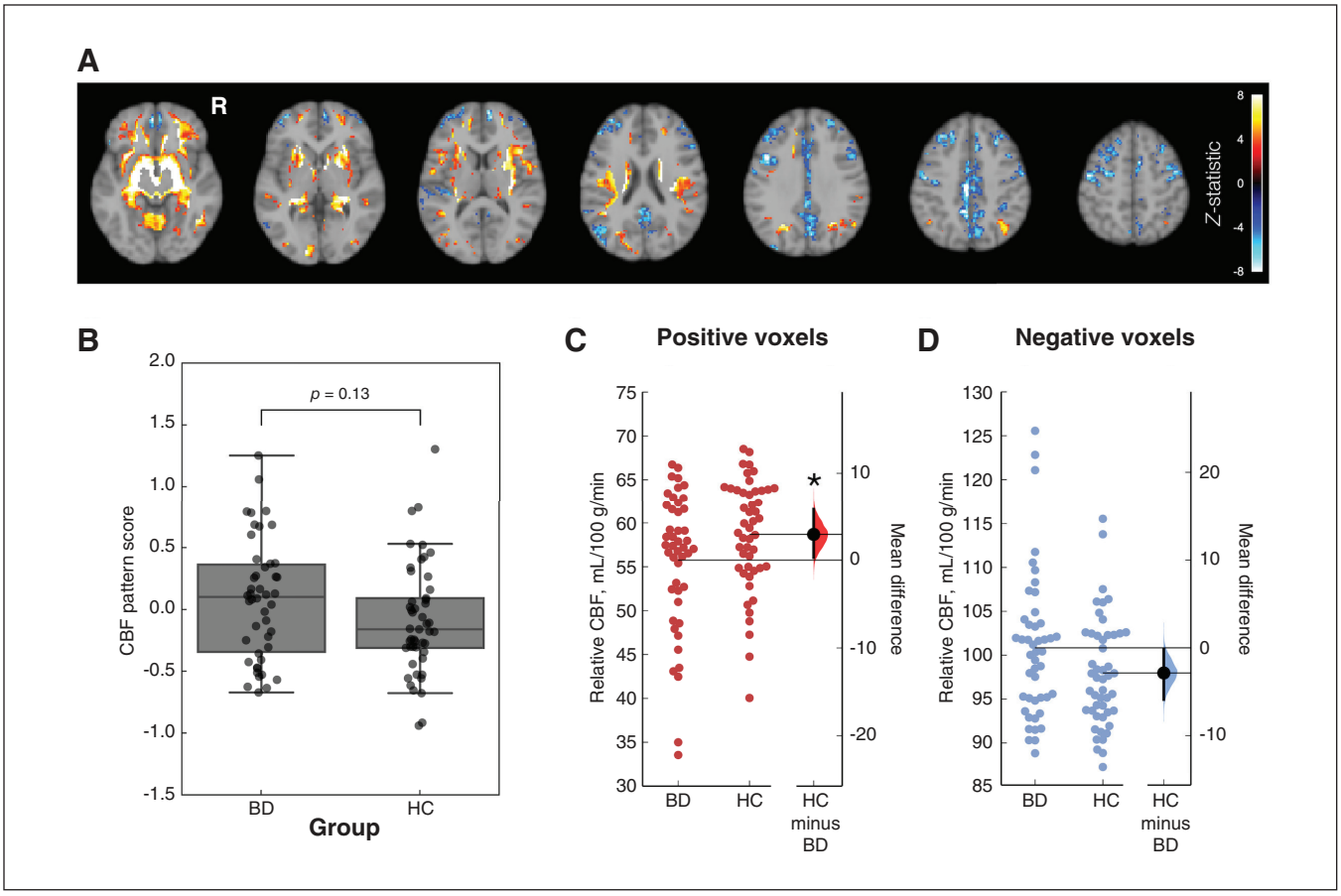

Figure 2: (A) Composite cerebral blood flow (CBF) pattern associated with bipolar disorder (BD), overlaid on a reference T_1-weighted image. Colours indicate *z*-scored voxel intensity and are unitless. No-coloured regions were not statistically significant after cluster-based thresholding. (B) Expression scores of the CBF pattern in the CBF image of each participant. We calculated a group comparison *p* value through the permutation-based null distribution of the difference in mean scores. Panels (C) and (D) show the mean relative CBF (intensity-normalized to a mean of 80 mL/100 g/min) and difference between youths with and without BD in the region-of-interest analysis for positive and negative voxels, respectively. Horizontal lines indicate group means. Error bars indicate the 95% confidence intervals. Density plots show the bootstrapped distribution and confidence intervals for the HCs minus the BD condition. HC = healthy control; R = right hemisphere. *$p < 0.05$.

factors may subserve the group differences, reflected in the higher cerebral metabolic rate of oxygen consumption and preliminary data showing higher brain temperature among adolescents with BD compared with controls.[49,50] On the other hand, numerous lines of investigation posit cellular and molecular mechanisms that could contribute to CBF differences, including cellular oxidative stress and metabolism, vascular dysfunction and mitochondria function.[47,48]

We observed a significant within-BD association between mean grey matter CBF and current depression score. This is consistent with a recent study that found CBF reductions in the right lateral occipital, angular and middle temporal gyrus in relation to depression score.[38] The current study also showed no association between CBF and mania score. We identified the principal components relevant to classification using forward stepwise regression, and used the AIC to assess goodness-of-fit. In this way, we attempted to limit the number of model features to those principal components with higher explained variance. Although disease-related variability may be explained by principal components accounting for a small fraction of total variance, the present approach was parsimonious. This approach avoided identifying spurious relationships between diagnostic group and sample-specific variability (i.e., it offered regularization capacity) as the AIC values were moderately variable across repeated cross-validation.

By combining principal components according to the regression coefficients of the logistic regression, we obtained a single brain map, weighted by the principal components' relevance in the classification problem. We used region-of-interest analysis to study the group-related CBF differences underlying this pattern. The composite CBF pattern identified regions with higher CBF in BD, consistent with the existing literature, although this difference was not significant. These regions included the cingulate gyrus and bilateral frontal pole, which have been previously identified in studies of CBF in youth BD; these regions contribute to reward processing in the brain.[14,15,39] The composite pattern also showed a set of regions with significantly lower relative CBF among youth with BD, compared with healthy controls. These included the insular gyrus, frontal gyrus and the ventral and dorsal striatum, which are reward- and mood-related regions that are thought to be central to the neurobiology of BD. In particular, both functional and structural differences have been found in these regions among youth with BD, compared with controls.[6,40,41] These findings of lower CBF did not emerge from previous studies and may be attributed to our multivariate analysis approach that yielded additional insight on the data. In contrast, previous studies used univariate analysis. We also used relative CBF in the region-of-interest analysis, whereas previous studies focused on quantitative CBF comparisons. We found decreased striatum CBF in the BD group; this finding provides some physiologic context for clinical and cognitive symptoms in BD, which may be related to altered dopamine signalling.[41] Neuroimaging readouts such as brain structure, regional brain activation and functional connectivity have been used to classify adult patients with BD versus healthy controls or adults with other psychiatric disorders.[42] One study showed that using MRI structure images and functional connectivity within the default-mode network and motor network had good performance in classifying BD subtypes and the number of episodes among adults with BD.[42] However, there is a limited understanding of classifying adolescents with and without BD and using CBF as a marker. Our work thus represents an important step toward classifying different clinical samples and adds to the current knowledge of youth BD classification.

The fALFF maps did not meaningfully classify youth with BD and healthy controls in the current study. This finding is at odds with a meta-analysis showing altered fMRI features in BD and another study showing good performance in classifying adult BD subtypes using functional connectivity data.[42,45] Differences in the fMRI hemodynamic response have also been observed when comparing BD and schizophrenia groups.[46] These findings suggest that fMRI-based approaches can classify adults with or without BD, but we used an exclusively adolescent sample. One interpretation of the current findings is that CBF data appear to be better suited to distinguish the 2 adolescent groups relative to the fALFF maps. More research on classification approaches and MRI-based data sources is warranted. Given the relative availability of fMRI data, it would be desirable to revisit the fMRI classification of adolescents with BD with a larger sample size.[44]

Limitations

It is important to note that the current study contrasted youth with BD with a nonclinical control group. A natural extension of this research would include additional clinical samples, for instance, to classify youth with BD, youth with unipolar depression and healthy controls. We cannot comment on which principal component CBF features are most relevant to the classification. We did not consider various combinations of these features beyond the ordering defined by explained variance. Notably, participants with BD were in various mood states and had various comorbidities, and most were taking medications, which may lead to challenges of small sample sizes and an unbalanced classification problem. In addition, we cannot rule out that group demographic differences were not a factor, despite adjusting for age. The univariate approach in which we used global grey matter CBF for classification had the advantage of being relatively easy to implement. Multivariate methods have the advantage of incorporating CBF data across all brain regions of interest; a natural extension from the current study would be to include additional clinical mood data or incorporate additional clinical samples (e.g., unipolar depression). The ASL-based CBF imaging had some limitations. This technique affords moderate test–retest repeatability;[43] however, youth tend to have high perfusion levels relative to adults, resulting in high-quality ASL images. Future studies could address intermediate psychiatric phenotypes, namely youth with unipolar depression.

Conclusion

We applied multivariate CBF features to distinguish youth with and without BD. We showed that CBF enabled classification with above-chance performance but is still suboptimal for clinical applications. A multivariate approach allowed exploration and explanation of the imaging features contributing to classification and found sets of regions with anomalous CBF that have been previously implicated in BD. Altogether, ASL MRI provides a vascular imaging readout that can be used to classify participant groups using a multivariate framework. We envision the inclusion of CBF measurements in future experiments aimed at developing quantitative metrics to elucidate BD symptoms, course of illness and treatment response.

Acknowledgements: The authors thank the youth and their parents for participating in this study, as well as the staff of the Centre for Youth Bipolar Disorder at the Centre for Addiction and Mental Health in Toronto.

Affiliations: From Hurvitz Brain Sciences, Sunnybrook Research Institute, Toronto, Ont. (Luciw, Jiang, Graham, MacIntosh); the Department of Medical Biophysics, University of Toronto, Toronto, Ont. (Luciw, Jiang, Chen, Graham, MacIntosh); the Centre for Youth Bipolar Disorder, Centre for Addiction and Mental Health, Toronto, Ont. (Grigorian, Dimick, Goldstein); the Department of Pharmacology and Toxicology, University of Toronto, Toronto, Ont. (Dimick, Goldstein); the Rotman Research Institute, Baycrest Health Sciences, Toronto, Ont. (Chen); the Institute of Biomedical Engineering, University of Toronto, Toronto, Ont. (Chen); the Department of Psychiatry, University of Toronto, Toronto, Ont. (Goldstein); the Sandra Black Centre for Brain Resilience & Recovery, Toronto, Ont. (MacIntosh); the Computational Radiology & Artificial Intelligence Unit, Oslo University Hospital, Norway (MacIntosh).

Competing interests: Bradley MacIntosh reports funding from the Brain and Behavior Research Foundation. Benjamin Goldstein is chair of the Pediatric Task Force and Vascular Task Force with the International Society for Bipolar Disorders. He holds the RBC Investments Chair in children's mental health and developmental psychopathology. No other competing interests were declared.

Funding: This study was funded by the Canadian Institutes of Health Research (no. 136947), the Natural Sciences and Engineering Research Council and the Ontario Mental Health Foundation (no. 1010589). Mikaela Dimick was supported by a CIHR Postdoctoral Award.

References

1. Grande I, Berk M, Birmaher B, et al. Bipolar disorder. *Lancet* 2016;387:1561-72.
2. Regier DA, Narrow WE, Clarke DE, et al. DSM-5 field trials in the United States and Canada, Part II: test-retest reliability of selected categorical diagnoses. *Am J Psychiatry* 2013;170:59-70.
3. Goetz M, Novak T, Vesela M, et al. Early stages of pediatric bipolar disorder: Retrospective analysis of a Czech impatient sample. *Neuropsychiatr Dis Treat* 2015;11:2855-64.
4. Vieta E, Salagre E, Grande I, et al. Early Intervention in bipolar disorder. *Am J Psychiatry* 2018;175:411-26.
5. Duffy A, Alda M, Hajek T, et al. Early stages in the development of bipolar disorder. *J Affect Disord* 2010;121:127-35.
6. Hibar DP, Westlye LT, Doan NT, et al. Cortical abnormalities in bipolar disorder: an MRI analysis of 6503 individuals from the ENIGMA Bipolar Disorder Working Group. *Mol Psychiatry* 2018;23:932-42.
7. Vargas C, López-Jaramillo C, Vieta E. A systematic literature review of resting state network — functional MRI in bipolar disorder. *J Affect Disord* 2013;150:727-35.
8. Toma S, MacIntosh BJ, Swardfager W, et al. Cerebral blood flow in bipolar disorder: a systematic review. *J Affect Disord* 2018;241:505-13.
9. Schnack HG, Nieuwenhuis M, van Haren NEM, et al. Can structural MRI aid in clinical classification? A machine learning study in two independent samples of patients with schizophrenia, bipolar disorder and healthy subjects. *Neuroimage* 2014;84:299-306.
10. Rocha-Rego V, Jogia J, Marquand AF, et al. Examination of the predictive value of structural magnetic resonance scans in bipolar disorder: a pattern classification approach. *Psychol Med* 2014;44:519-32.
11. Nunes A, Schnack HG, Ching CRK, et al. Using structural MRI to identify bipolar disorders — 13 site machine learning study in 3020 individuals from the ENIGMA Bipolar Disorders Working Group. *Mol Psychiatry* 2020;25:2130-43.
12. Almeida JRC, Mourao-Miranda J, Aizenstein HJ, et al. Pattern recognition analysis of anterior cingulate cortex blood flow to classify depression polarity. *Br J Psychiatry* 2013;203:310-1.
13. Goldstein BI, Carnethon MR, Matthews KA, et al. Major depressive disorder and bipolar disorder predispose youth to accelerated atherosclerosis and early cardiovascular disease: a scientific statement from the American Heart Association. *Circulation* 2015;132:965-86.
14. MacIntosh BJ, Shirzadi Z, Scavone A, et al. Increased cerebral blood flow among adolescents with bipolar disorder at rest is reduced following acute aerobic exercise. *J Affect Disord* 2017;208:205-13.
15. Karthikeyan S, Fiksenbaum L, Grigorian A, et al. Normal cerebral oxygen consumption despite elevated cerebral blood flow in adolescents with bipolar disorder: putative neuroimaging evidence of anomalous energy metabolism. *Front Psychiatry* 2019;10:739.
16. Dimick MK, Toma S, MacIntosh BJ, et al. Cerebral blood flow and core mood symptoms in youth bipolar disorder: evidence for region-symptom specificity. *J Am Acad Child Adolesc Psychiatry* 2022;61:1455-65.
17. Moeller JR, Strother SC, Sidtis JJ, et al. Scaled subprofile model: a statistical approach to the analysis of functional patterns in positron emission tomographic data. *J Cereb Blood Flow Metab* 1987;7:649-58.
18. Melzer TR, Watts R, MacAskill MR, et al. Arterial spin labelling reveals an abnormal cerebral perfusion pattern in Parkinson's disease. *Brain* 2011;134:845-55.
19. Kaufman J, Birmaher B, Brent D, et al. Schedule for affective disorders and schizophrenia for school-age children-present and lifetime version (K-SADS-PL): initial reliability and validity data. *J Am Acad Child Adolesc Psychiatry* 1997;36:980-8.
20. Chambers WJ, Puig-Antich J, Hirsch M, et al. The assessment of affective disorders in children and adolescents by semistructured interview: test-retest reliability of the schedule for affective disorders and schizophrenia for school-age children, present episode version. *Arch Gen Psychiatry* 1985;42:696-702.
21. Axelson D, Birmaher BJ, Brent D, et al. A preliminary study of the kiddie schedule for affective disorders and schizophrenia for school-age children mania rating scale for children and adolescents. *J Child Adolesc Psychopharmacol* 2003;13:463-70.
22. Axelson D, Birmaher B, Strober M, et al. Phenomenology of children and adolescents with bipolar spectrum disorders. *Arch Gen Psychiatry* 2006;63:1139-48.

23. Weissman MM, Wickramaratne P, Adams P, et al. Brief screening for family psychiatric history: the Family History Screen. *Arch Gen Psychiatry* 2000;57:675-82.
24. Shaffer D, Gould MS, Brasic J, et al. A Children's Global Assessment Scale (CGAS). *Arch Gen Psychiatry* 1983;40:1228-31.
25. *National Health and Nutrition Examination Survey (NHANES) anthropometry procedures manual*. Hyattsville (MD): Department of Health and Human Services, Centers for Disease Control and Prevention; 2007.
26. Smith SM, Jenkinson M, Woolrich MW, et al. Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage* 2004;23:S208-19.
27. Liu TT, Wong EC. A signal processing model for arterial spin labeling functional MRI. *Neuroimage* 2005;24:207-15.
28. Shirzadi Z, Stefanovic B, Chappell MA, et al. Enhancement of automated blood flow estimates (ENABLE) from arterial spin-labeled MRI. *J Magn Reson Imaging* 2018;47:647-55.
29. Alsop DC, Detre JA, Golay X, et al. Recommended implementation of arterial spin-labeled Perfusion MRI for clinical applications: a consensus of the ISMRM Perfusion Study Group and the European Consortium for ASL in Dementia. *Magn Reson Med* 2015;73:102-16.
30. Asllani I, Borogovac A, Brown TR. Regression algorithm correcting for partial volume effects in arterial spin labeling MRI. *Magn Reson Med* 2008;60:1362-71.
31. Mutsaerts HJMM, Petr J, Groot P, et al. ExploreASL: an image processing pipeline for multi-center ASL perfusion MRI studies. *Neuroimage* 2020;219:117031.
32. Zhang Y, Brady M, Smith S. Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. *IEEE Trans Med Imaging* 2001;20:45-57.
33. Zou Q-H, Zhu C-Z, Yang Y, et al. An improved approach to detection of amplitude of low-frequency fluctuation (ALFF) for resting-state fMRI: Fractional ALFF. *J Neurosci Methods* 2008;172:137-41.
34. Wu T, Ma Y, Zheng Z, et al. Parkinson's disease-related spatial covariance pattern identified with resting-state functional MRI. *J Cereb Blood Flow Metab* 2015;35:1764-70.
35. Akaike H. A new look at the statistical model identification. *IEEE Trans Automat Contr* 1974;19:716-23.
36. Woo C-W, Krishnan A, Wager TD. Cluster-extent based thresholding in fMRI analyses: pitfalls and recommendations. *Neuroimage* 2014;91:412-9.
37. Spetsieris PG, Eidelberg D. Scaled subprofile modeling of resting state imaging data in Parkinson's disease: methodological issues. *Neuroimage* 2011;54:2899-914.
38. Zeng V, Lizano P, Bolo NR, et al. Altered cerebral perfusion in bipolar disorder: A pCASL MRI study. *Bipolar Disord* 2021;23:130-40.
39. Toma S, MacIntosh B, Grigorian A, et al. F133. Cerebral blood flow is altered according to mood states in adolescents with bipolar disorder. *Biol Psychiatry* 2019;85:S265.
40. Phillips ML, Swartz HA. A critical appraisal of neuroimaging studies of bipolar disorder: toward a new conceptualization of underlying neural circuitry and a road map for future research. *Am J Psychiatry* 2014;171:829-43.
41. Leibenluft E, Charney DS, Pine DS. Researching the pathophysiology of pediatric bipolar disorder. *Biol Psychiatry* 2003;53:1009-20.
42. Chen YL, Tu P, Huang T, et al. Identifying subtypes of bipolar disorder based on clinical and neurobiological characteristics. *Sci Rep* 2021;11:17082.
43. Chen Y, Wang DJJ, Detre JA. Test-retest reliability of arterial spin labeling with common labeling strategies. *J Magn Reson Imaging* 2011;33:940-9.
44. Hernandez-Garcia L, Aramendía-Vidaurreta V, Bolar DS, et al. Recent technical developments in ASL: a review of the state of the art. *Magn Reson Med* 2022;88:2021-42.
45. Yoon S, Kim TD, Kim J, et al. Altered functional activity in bipolar disorder: A comprehensive review from a large-scale network perspective. *Brain Behav* 2021;11:e01953.
46. Yan W, Palaniyappan L, Liddle PF, et al. Characterization of hemodynamic alterations in schizophrenia and bipolar disorder and their effect on resting-state fMRI functional connectivity. *Schizophr Bull* 2022;48:695-711.
47. Andreazza AC, Duong A, Young LT. Bipolar disorder as a mitochondrial disease. *Biol Psychiatry* 2018;83:720-1.
48. Daiber A, Steven S, Weber A, et al. Targeting vascular (endothelial) dysfunction. *Br J Pharmacol* 2017;174:1591-619.
49. Karthikeyan S, Fiksenbaum L, Grigorian A, et al. Normal cerebral oxygen consumption despite elevated cerebral blood flow in adolescents with bipolar disorder: putative neuroimaging evidence of anomalous energy metabolism. *Front Psychiatry* 2019;10:739.
50. Zou Y, Heyn C, Grigorian A, et al. Measuring brain temperature in youth bipolar disorder using a novel magnetic resonance imaging approach: a proof-of-concept study. *Curr Neuropharmacol* 2023;21:1355-66.