https://doi.org/10.70401/cbm.2026.0009

Supplementary information

PCMCI-SVM: A model identifying diagnostic biomarkers for autism spectrum disorder through causal network analysis

Hao Wang, Lingzhi Liu, Yanting Liu, Yanrui Ding

Table S1. AAL atlas regions along with corresponding ROI names. In the second column, the R at the end of the brain region name indicates right and the L indicates left.

Serial number	Name of brain region (abbreviated)	Name of brain region
2001	PreCG.L	Precentral gyrus
2002	PreCG.R	Precentral gyrus
2101	SFGdor.L	Superior frontal gyrus, dorsolateral
2102	SFGdor.R	Superior frontal gyrus, dorsolateral
2111	ORBsup.L	Superior frontal gyrus, orbital part
2112	ORBsup.R	Superior frontal gyrus, orbital part
2201	MFG.L	Middle frontal gyrus
2202	MFG.R	Middle frontal gyrus
2211	ORBmid.L	Middle frontal gyrus, orbital part
2212	ORBmid.R	Middle frontal gyrus, orbital part
2301	IFGoperc.L	Inferior frontal gyrus, opercular part
2302	IFGoperc.R	Inferior frontal gyrus, opercular part
2311	IFGtriang.L	Inferior frontal gyrus, triangular part
2312	IFGtriang.R	Inferior frontal gyrus, triangular part
2321	ORBinf.L	Inferior frontal gyrus, orbital part
2322	ORBinf.R	Inferior frontal gyrus, orbital part
2331	ROL.L	Rolandic operculum
2332	ROL.R	Rolandic operculum
2401	SMA.L	Supplementary motor area
2402	SMA.R	Supplementary motor area
2501	OLF.L	Olfactory cortex
2502	OLF.R	Olfactory cortex
2601	SFGmed.L	Superior frontal gyrus, medial
2602	SFGmed.R	Superior frontal gyrus, medial
2611	ORBsupmed.L	Superior frontal gyrus, medial orbital
2612	ORBsupmed.R	Superior frontal gyrus, medial orbital
2701	REC.L	Gyrus rectus
2702	REC.R	Gyrus rectus
3001	INS.L	Insula
3002	INS.R	Insula
4001	ACG.L	Anterior cingulate and paracingulate gyri

4002	ACG.R	Anterior cingulate and paracingulate gyri
4011	DCG.L	Median cingulate and paracingulate gyri
4012	DCG.R	Median cingulate and paracingulate gyri
4021	PCG.L	Posterior cingulate gyrus
4022	PCG.R	Posterior cingulate gyrus
4101	HIP.L	Hippocampus
4102	HIP.R	Hippocampus
4111	PHG.L	Parahippocampal gyrus
4112	PHG.R	Parahippocampal gyrus
4201	AMYG.L	Amygdala
4202	AMYG.R	Amygdala
5001	CAL.L	Calcarine fissure and surrounding cortex
5002	CAL.R	Calcarine fissure and surrounding cortex
5011	CUN.L	Cuneus
5012	CUN.R	Cuneus
5021	LING.L	Lingual gyrus
5022	LING.R	Lingual gyrus
5101	SOG.L	Superior occipital gyrus
5102	SOG.R	Superior occipital gyrus
5201	MOG.L	Middle occipital gyrus
5202	MOG.R	Middle occipital gyrus
5301	IOG.L	Inferior occipital gyrus
5302	IOG.R	Inferior occipital gyrus
5401	FFG.L	Fusiform gyrus
5402	FFG.R	Fusiform gyrus
6001	PoCG.L	Postcentral gyrus
6002	PoCG.R	Postcentral gyrus
6101	SPG.L	Superior parietal gyrus
6102	SPG.R	Superior parietal gyrus
6201	IPL.L	Inferior parietal, but supramarginal and angular gyri
6202	IPL.R	Inferior parietal, but supramarginal and angular gyri
6211	SMG.L	Supramarginal gyrus
6212	SMG.R	Supramarginal gyrus

6221	ANG.L	Angular gyrus
6222	ANG.R	Angular gyrus
6301	PCUN.L	Precuneus
6302	PCUN.R	Precuneus
6401	PCL.L	Paracentral lobule
6402	PCL.R	Paracentral lobule
7001	CAU.L	Caudate nucleus
7002	CAU.R	Caudate nucleus
7011	PUT.L	Lenticular nucleus, putamen
7012	PUT.R	Lenticular nucleus, putamen
7021	PAL.L	Lenticular nucleus, pallidum
7022	PAL.R	Lenticular nucleus, pallidum
7101	THA.L	Thalamus
7102	THA.R	Thalamus
8101	HES.L	Heschl gyrus
8102	HES.R	Heschl gyrus
8111	STG.L	Superior temporal gyrus
8112	STG.R	Superior temporal gyrus
8121	TPOsup.L	Temporal pole: superior temporal gyrus
8122	TPOsup.R	Temporal pole: superior temporal gyrus
8201	MTG.L	Middle temporal gyrus
8202	MTG.R	Middle temporal gyrus
8211	TPOmid.L	Temporal pole: middle temporal gyrus
8212	TPOmid.R	Temporal pole: middle temporal gyrus
8301	ITG.L	Inferior temporal gyrus
8302	ITG.R	Inferior temporal gyrus
9001	CRBLCrus1.L	Cerebellum_Superior
9002	CRBLCrus1.R	Cerebellum_Superior
9011	CRBLCrus2.L	Cerebellum_Inferior
9012	CRBLCrus2.R	Cerebellum_Inferior
9021	CRBL3.L	Cerebellum_Superior
9022	CRBL3.R	Cerebellum_Superior
9031	CRBL45.L	Cerebellum_Superior
9032	CRBL45.R	Cerebellum_Superior
9041	CRBL6.L	Cerebellum_Superior

9042	CRBL6.R	Cerebellum_Superior
9051	CRBL7b.L	Cerebellum_Inferior
9052	CRBL7b.R	Cerebellum_Inferior
9061	CRBL8.L	Cerebellum_Inferior
9062	CRBL8.R	Cerebellum_Inferior
9071	CRBL9.L	Cerebellum_Inferior
9072	CRBL9.R	Cerebellum_Inferior
9081	CRBL10.L	Cerebellum_Inferior
9082	CRBL10.R	Cerebellum_Inferior
9100	Vermis12	Vermis
9110	Vermis3	Vermis
9120	Vermis45	Vermis
9130	Vermis6	Vermis
9140	Vermis7	Vermis
9150	Vermis8	Vermis
9160	Vermis9	Vermis
9170	Vermis_10	Vermis10

ROI: region of interest.

Based on the results of the two-sample t-test p-value, we obtained the 30 most significant causal connections for each baseline method: spectral dynamic causal modeling (spDCM; Table S2), Granger causal analysis (Table S3), transfer entropy (TE; Table S4), and Liang-Kleeman information flow (Table S5). These tables display group-averaged alterations in causal connection strength between autism spectrum disorder (ASD) and typical control (TC). The results in these tables provide strong support for our in-depth understanding of causal connectivity differences between ASD and TC.

Table S2. Causal connectivity difference in spDCM causal network.

Method	Brain region 1	Brain region 2	P-value	Causal connection change
DCM	SPG.R	HES.R	7.49E-06	decrease
DCM	PCUN.R	TPOmid.R	0.000114	decrease
DCM	IOG.L	Vermis6	0.000165	increase
DCM	PCG.L	CRBL45.L	0.000237	increase
DCM	HIP.R	DCG.R	0.000348	increase
DCM	PUT.R	IFGtriang.R	0.00048	increase

DCM	CUN.R	DCG.L	0.000499	increase
DCM	SMG.L	SMG.R	0.000512	decrease
DCM	ITG.L	Vermis8	0.000758	decrease
DCM	IFGtriang.R	MFG.L	0.000796	increase
DCM	ANG.L	PCUN.R	0.000876	increase
DCM	HIP.L	PHG.R	0.001114	decrease
DCM	CRBLCrus1.L	SOG.R	0.00119	decrease
DCM	PoCG.L	SMG.L	0.001395	decrease
DCM	IFGoperc.L	ROL.L	0.001482	increase
DCM	IPL.R	PreCG.L	0.001759	decrease
DCM	SPG.R	CUN.R	0.001971	decrease
DCM	SMG.R	FFG.L	0.002096	decrease
DCM	CRBL45.L	SMG.R	0.00213	decrease
DCM	Vermis9	CRBL6.R	0.002166	increase
DCM	MFG.R	HIP.R	0.002258	decrease
DCM	CAL.R	MFG.R	0.002516	increase
DCM	ACG.L	CRBL7b.L	0.002565	increase
DCM	SPG.L	SFGdor.R	0.002615	decrease
DCM	MFG.L	PoCG.L	0.003001	increase
DCM	OLF.R	CRBL8.R	0.003114	increase
DCM	SMA.L	AMYG.R	0.003231	decrease
DCM	HES.L	SMG.R	0.00327	decrease
DCM	ORBsupmed.R	IPL.L	0.003331	increase
DCM	IPL.R	SMA.L	0.00343	decrease

The direction of the causal connection is: Brain region 1 → Brain region 2.

Table S3. Causal connectivity difference in Granger causal network.

Method	Brain region 1	Brain region 2	P-value	Causal connection change
Granger	CRBL7b.R	PUT.L	2.74E-05	decrease
Granger	HES.R	LING.L	0.000144	decrease
Granger	SMA.R	Vermis3	0.000267	decrease
Granger	ORBsupmed.L	MTG.L	0.000388	decrease
Granger	ORBsup.L	CRBL45.L	0.000526	decrease
Granger	PUT.R	IPL.R	0.000549	decrease
Granger	FFG.R	ANG.L	0.000615	increase

Granger	IFGtriang.L	IFGoperc.L	0.000641	decrease
Granger	ITG.R	IOG.R	0.000687	decrease
Granger	ROL.L	SOG.L	0.000797	increase
Granger	IPL.L	LING.L	0.00081	decrease
Granger	MTG.L	PUT.R	0.000904	decrease
Granger	MTG.L	Vermis6	0.000909	decrease
Granger	Vermis9	ITG.R	0.000938	decrease
Granger	ORBsupmed.R	SMG.L	0.00094	decrease
Granger	IPL.L	CRBL3.R	0.000979	decrease
Granger	CRBL3.R	IPL.L	0.0012	decrease
Granger	ACG.R	CRBLCrus1.L	0.001297	decrease
Granger	STG.L	CRBLCrus1.R	0.001379	decrease
Granger	PHG.L	CRBL10.R	0.001385	decrease
Granger	FFG.R	MTG.R	0.001522	decrease
Granger	PUT.L	CRBL7b.R	0.001539	decrease
Granger	STG.L	CRBLCrus2.L	0.00158	decrease
Granger	SFGmed.R	PCUN.L	0.001632	decrease
Granger	CRBL3.L	LING.L	0.001665	increase
Granger	FFG.R	IFGoperc.R	0.001773	increase
Granger	HES.R	PCG.L	0.00198	increase
Granger	ORBsupmed.L	ORBinf.R	0.002079	decrease
Granger	TPOsup.R	CUN.L	0.002096	increase
Granger	Vermis45	SFGdor.L	0.002239	decrease

The direction of the causal connection is: Brain region 1 → Brain region 2.

Table S4. Causal connectivity difference in TE causal network.

Method	Brain region 1	Brain region 2	P-value	Causal connection change
TE	FFG.R	CAL.L	1.76E-05	decrease
TE	INS.R	SMA.R	7.40E-05	decrease
TE	ORBinf.L	PCUN.L	7.84E-05	decrease
TE	SFGmed.R	ITG.R	0.00026	increase
TE	ORBsup.L	SMA.L	0.00029	decrease
TE	CRBLCrus2.R	Vermis3	0.000309	decrease
TE	PCUN.R	TPOmid.L	0.000421	increase
TE	CAU.R	LING.L	0.000597	decrease

TE	CAU.R	CAL.L	0.00064	decrease
TE	LING.L	IFGtriang.R	0.000982	increase
TE	TPOmid.L	STG.L	0.001054	increase
TE	PCL.R	CRBLCrus1.L	0.001122	decrease
TE	PoCG.R	CRBLCrus2.L	0.001125	decrease
TE	CRBL6.L	MTG.L	0.001265	increase
TE	ORBsup.L	LING.R	0.001331	decrease
TE	PHG.L	CRBLCrus1.L	0.001472	decrease
TE	PUT.L	Vermis10	0.001663	increase
TE	SFGmed.L	ITG.R	0.00168	increase
TE	SMG.R	INS.L	0.001811	decrease
TE	OLF.L	ITG.L	0.001821	decrease
TE	PUT.L	PreCG.L	0.001844	decrease
TE	HIP.R	ORBsup.R	0.001908	decrease
TE	INS.L	CAU.L	0.001951	increase
TE	CAU.R	CRBL7b.L	0.001962	decrease
TE	PreCG.L	Vermis3	0.001989	decrease
TE	ORBsupmed.R	AMYG.L	0.002282	increase
TE	CAU.L	PHG.R	0.002282	increase
TE	PCL.R	CRBLCrus2.R	0.002346	decrease
TE	PAL.L	SFGmed.R	0.00239	increase
TE	ORBmid.L	DCG.L	0.002414	decrease

The direction of the causal connection is: Brain region 1 → Brain region 2.

Table S5. Causal connectivity difference in Liang-Kleeman information flow causal network.

Method	Brain region 1	Brain region 2	P-value	Causal connection change
Liang	CRBL6.R	Vermis45	3.01E-05	increase
Liang	CRBL6.R	MOG.R	3.76E-05	increase
Liang	PreCG.R	MTG.R	0.000151	increase
Liang	CRBL6.R	SOG.R	0.000156	increase
Liang	CRBL3.L	ROL.R	0.000365	increase
Liang	ANG.L	ORBinf.L	0.000393	decrease
Liang	MFG.R	Vermis3	0.000532	decrease
Liang	PHG.R	CRBLCrus2.L	0.000574	decrease
Liang	CRBL6.R	LING.R	0.000598	increase

Liang	IFGoperc.L	SPG.L	0.000735	decrease
Liang	CRBL10.L	Vermis10	0.00086	increase
Liang	AMYG.R	PCUN.L	0.000926	decrease
Liang	THA.R	MOG.R	0.00099	decrease
Liang	CRBL6.R	SFGdor.R	0.001214	increase
Liang	TPOmid.L	ROL.R	0.001317	increase
Liang	PUT.L	ORBsup.R	0.001384	increase
Liang	ROL.L	IOG.L	0.001486	increase
Liang	SMG.L	TPOsup.L	0.001894	decrease
Liang	CAL.L	ORBsupmed.L	0.001996	increase
Liang	Vermis3	CRBL10.L	0.002066	increase
Liang	PUT.L	AMYG.R	0.002355	decrease
Liang	CRBL6.R	SFGdor.L	0.002371	increase
Liang	CUN.R	PCUN.L	0.002767	increase
Liang	Vermis10	ROL.L	0.002783	decrease
Liang	SMG.R	TPOsup.L	0.002799	increase
Liang	ANG.L	SFGmed.L	0.002829	decrease
Liang	CRBL9.L	OLF.R	0.002865	decrease
Liang	LING.R	PUT.L	0.003046	decrease
Liang	CAU.L	PCL.R	0.003071	increase
Liang	CRBL9.R	CRBL3.L	0.003375	increase

The direction of the causal connection is: Brain region 1 → Brain region 2.

Two-sample t-tests assess differences in causal connections between ASD and TC. The significance level for the two-sample t-test is set at 0.05 to find significant differences in causal connectivity. Based on causal connectivity with significant difference, we attempt to identify ASD from a potentially at-risk population. Causal connectivity with significant difference is used as characteristics of each participant (The features chosen are the same as those used for SVM classification). Then, we use RF and KNN for classification and the results are shown in Table S6 and Table S7. A grid search is conducted to select the optimal parameters for the RF and KNN models. In the RF classifier, parameters are optimized using a grid search method to balance model performance and computational efficiency. Parameters such as n_estimators, max_depth, and min_samples_split are systematically adjusted to find the optimal model configuration. For the KNN classifier, a grid search strategy is also employed to optimize parameters like n_neighbors, weights, and metric, aiming to achieve the best classification results.

Table S6. The classification results of ASD based on RF using different causal models.

Model	Test-ACC	SEN	SPE	F1-Score	AUC	Train-ACC	Validation-ACC
PCMCI-RF	80.94%	85.03%	76.43%	0.8305	0.9096	99.86%	84.77%
spDCM-RF	75.18%	88.72%	58.59%	0.798	0.8513	99.74%	77.34%
Granger-RF	85.29%	90.69%	79.43%	0.8718	0.936	98.98%	82.51%
TE-RF	80.24%	85.18%	74.84%	0.8252	0.8663	99.81%	82.76%
LKIF-RF	75.18%	87.49%	60.14%	0.7946	0.8031	99.86%	77.19%

Test-ACC/Train-ACC/Validation-ACC: the average accuracy of 25 replicate experiments.

Table S7. The classification results of ASD based on KNN using different causal models.

Model	Test-ACC	SEN	SPE	F1-Score	AUC	Train-ACC	Validation-ACC
PCMCI-KNN	86.59%	86.18%	87.99%	0.8777	0.9392	93.50%	88.92%
spDCM-KNN	77.41%	86.74%	65.50%	0.8104	0.8598	86.45%	79.85%
Granger-KNN	75.88%	57.23%	99.25%	0.7224	0.9007	85.48%	76.88%
TE-KNN	76.47%	66.00%	90.03%	0.7549	0.8663	88.02%	83.55%
LKIF-KNN	70.71%	93.00%	41.97%	0.7783	0.7594	82.71%	71.45%

Test-ACC/Train-ACC/Validation-ACC: the average accuracy of 25 replicate experiments.

In constructing ASD-TC classification model based on RF and KNN algorithm, the test set constitutes 20% of the total dataset and is used to evaluate the generalization ability of the model. In dividing the test set, five random seeds 42-46 are used. To evaluate the performance of the model, we perform ten-fold cross-validation. The validation set accounts for 8% of the total dataset and is used for tuning the model parameters as well as for preliminary assessment of the model capability. The training set constitutes 72% of the total dataset and is used for fitting model and updating model parameters. In dividing the validation set, five random seeds 42-46 are used because the results of ten-fold cross-validation are different for different random seeds. The optimal parameter combinations for SVM, KNN, and RF are respectively shown in Table S8,S9,S10. Using different random seeds for the division of the test and validation sets leads to varying results. Twenty-five unique outcomes are generated through the use of diverse random seed combinations, and the average accuracies are reported.

Table S8. Optimal parameters under different combinations of random seeds for the PCMCI-SVM model.

No.	Random seed for splitting the training and test sets	Random seed for 10-fold cross-validation	C	gamma	kernel	Test set accuracy
1	42	42	10	0.01	rbf	88.24%
2	42	43	10	0.01	rbf	88.24%
3	42	44	10	0.01	rbf	88.24%
4	42	45	10	scale	rbf	91.18%
5	42	46	10	auto	rbf	88.24%
6	43	42	10	auto	rbf	91.18%
7	43	43	1	0.1	rbf	97.06%
8	43	44	10	auto	rbf	91.18%
9	43	45	10	auto	rbf	91.18%
10	43	46	0.1	0.001	linear	97.06%
11	44	42	10	0.1	rbf	94.12%
12	44	43	1	scale	rbf	97.06%
13	44	44	10	scale	rbf	94.12%
14	44	45	10	scale	rbf	94.12%
15	44	46	10	scale	rbf	94.12%
16	45	42	100	0.001	rbf	85.29%
17	45	43	1	0.1	rbf	85.29%
18	45	44	1	0.1	rbf	85.29%
19	45	45	10	0.01	rbf	85.29%
20	45	46	10	0.01	rbf	85.29%
21	46	42	1	0.1	rbf	94.12%
22	46	43	10	0.01	rbf	91.18%
23	46	44	1	scale	rbf	97.06%
24	46	45	1	0.1	rbf	94.12%
25	46	46	1	0.1	rbf	94.12%

Table S9. Optimal parameters under different combinations of random seeds for the PCMCI-KNN model.

No.	Random seed for splitting the train and test sets	Random seed for 10-fold cross-validation	metric	n_neighbors	weights	Test set accuracy
1	42	42	manhattan	5	uniform	88.24%
2	42	43	manhattan	5	uniform	88.24%
3	42	44	manhattan	5	uniform	88.24%
4	42	45	manhattan	5	uniform	88.24%
5	42	46	manhattan	5	uniform	88.24%
6	43	42	manhattan	7	uniform	91.18%
7	43	43	manhattan	9	uniform	94.12%
8	43	44	manhattan	9	uniform	94.12%
9	43	45	manhattan	9	uniform	94.12%
10	43	46	manhattan	9	uniform	94.12%
11	44	42	manhattan	9	uniform	88.24%
12	44	43	manhattan	7	uniform	88.24%
13	44	44	manhattan	9	uniform	88.24%
14	44	45	manhattan	9	uniform	88.24%
15	44	46	manhattan	9	uniform	88.24%
16	45	42	manhattan	9	uniform	76.47%
17	45	43	manhattan	9	uniform	76.47%
18	45	44	manhattan	9	uniform	76.47%
19	45	45	manhattan	5	uniform	70.59%
20	45	46	manhattan	5	uniform	70.59%
21	46	42	euclidean	5	uniform	91.18%
22	46	43	euclidean	7	uniform	88.24%
23	46	44	euclidean	9	uniform	88.24%
24	46	45	euclidean	9	uniform	88.24%
25	46	46	euclidean	7	uniform	88.24%

Table S10. Optimal parameters under different combinations of random seeds for the PCMCI-RF model.

No.	Random seed for splitting the train and test sets	Random seed for 10-fold cross-validation	max depth	max features	min samples_leaf	min samples_split	n estimators	Test set accuracy
1	42	42	10	sqrt	1	5	50	73.53%
2	42	43	10	sqrt	4	5	100	79.41%
3	42	44	10	sqrt	2	5	50	79.41%
4	42	45	10	sqrt	1	5	100	76.47%
5	42	46	10	sqrt	2	5	100	82.35%
6	43	42	5	0.5	1	5	50	79.41%
7	43	43	10	sqrt	1	5	100	82.35%
8	43	44	7	sqrt	4	5	100	79.41%
9	43	45	10	sqrt	2	5	50	79.41%
10	43	46	10	sqrt	1	10	100	79.41%
11	44	42	7	sqrt	2	5	100	88.24%
12	44	43	7	sqrt	2	10	100	79.41%
13	44	44	10	sqrt	1	10	100	85.29%
14	44	45	7	sqrt	1	10	100	73.53%
15	44	46	10	sqrt	1	5	100	82.35%
16	45	42	10	sqrt	1	10	50	76.47%
17	45	43	7	sqrt	1	5	100	82.35%
18	45	44	7	sqrt	4	5	50	79.41%
19	45	45	10	sqrt	4	5	50	85.29%
20	45	46	10	sqrt	2	5	100	88.24%
21	46	42	10	sqrt	1	10	50	91.18%
22	46	43	7	sqrt	4	5	50	79.41%
23	46	44	5	sqrt	1	5	50	79.41%
24	46	45	10	sqrt	4	10	50	82.35%
25	46	46	10	sqrt	4	5	100	79.41%