

# SCIENTIFIC REPORTS



OPEN

## Identifying individuals with attention deficit hyperactivity disorder based on temporal variability of dynamic functional connectivity

Xun-Heng Wang<sup>1</sup>, Yun Jiao<sup>2</sup> & Lihua Li<sup>1</sup>

Attention deficit hyperactivity disorder (ADHD) is a common disorder that emerges in school-age children. The diagnostic model based on neuroimaging features could be beneficial for ADHD in twofold: identifying individuals with ADHD and discovering the discriminative patterns for patients. The dynamic functional connectivity of ADHD remains unclear. Towards this end, 100 children with ADHD and 140 normal controls were obtained from the ADHD-200 Consortium. The raw features were derived from the temporal variability between intrinsic connectivity networks (ICNs) as well as the demographic and covariate variables. The diagnostic model was based on the support vector machines (SVMs). The performance of diagnostic model was analyzed using leave-one-out cross-validation (LOOCV) and 10-folds cross-validations (CVs). The diagnostic model based on inter-ICN variability outperformed that based on inter-ICN functional connectivity and inter-ICN phase synchrony. The LOOCV achieved total accuracy of 78.75%, the sensitivity of 76%, and the specificity of 80.71%. The 10-folds CVs achieved average prediction accuracy of  $75.54\% \pm 1.34\%$ , average sensitivity of  $70.5\% \pm 2.34\%$ , and average specificity of  $77.44\% \pm 1.47\%$ . In addition, the discriminative patterns for ADHD were discovered using SVMs. The discriminative patterns confirmed with previous findings. In summary, individuals with ADHD could be identified through inter-ICN variability, which could be potential biomarkers for diagnostic model of ADHD.

Attention deficit hyperactivity disorder (ADHD) is a common disorder that spreads in school-age children<sup>1</sup>. According to epidemiological survey, ADHD affects nearly 5–10% of children and 4% of adults<sup>2</sup>. Patients with ADHD always exhibited problematic behaviors (i.e., inattention, impulsivity), academic failure, social dysfunction in their daily life<sup>3</sup>. Therefore, accurate diagnosis of ADHD is beneficial for individuals, as well as their related family and society. Clinical diagnostic models for ADHD were based on behavioral scales, which was subjective during implementation. The structural and functional MRI-based evidences suggested that the brain of ADHD might wire differently from healthy controls<sup>4–7</sup>, thus neuroimaging features might be potential biomarkers for diagnostic model of ADHD. Several neuroimaging-based diagnostic models have been established for ADHD using machine learning<sup>8–18</sup>. The diagnostic models could be beneficial for ADHD in twofold: classification of patients and discovery of ADHD-related discriminative patterns. However, it is a challenging task to classify individuals with ADHD from healthy controls based on brain imaging features. One challenge is the significant high dimensions of features in the diagnostic models. Another challenge is extracting novel features that could discriminate ADHD. On the one hand, the high dimensional features could increase the complexity of diagnostic models. On the other hand, novel features with low dimensions could be beneficial for diagnostic models of ADHD. So far, the dynamic functional connectivity of ADHD remain unexplored.

<sup>1</sup>College of Life Information Science and Instrument Engineering, Hangzhou Dianzi University, Hangzhou, 310018, China. <sup>2</sup>Jiangsu Key Laboratory of Molecular and Functional Imaging, Department of Radiology, Zhongda Hospital, Medical School of Southeast University, Nanjing, 210009, China. Correspondence and requests for materials should be addressed to X.-H.W. (email: [xhwang@hdu.edu.cn](mailto:xhwang@hdu.edu.cn)) or L.L. (email: [lilh@hdu.edu.cn](mailto:lilh@hdu.edu.cn))

performance	accuracy	sensitivity	specificity
VAR	78.75%	76%	80.71%
FC	72.92%	65%	78.57%
PS	68.75%	64%	72.14%

**Table 1.** Performance of the diagnostic model based on LOOCV.

Intrinsic connectivity networks (ICN) are spatially independent resting state networks, which are intrinsically dynamic and anti-correlated spontaneous brain systems<sup>19,20</sup>. There are about ten to twenty well-established ICNs found by independent component analysis (ICA)<sup>21–23</sup>. The ICNs were related to specific brain functions and even correlated to behavioral symptoms<sup>24</sup>. The ICNs were consistent across different individuals<sup>25</sup>. Thus, the ICNs might be potential biomarkers for brain disorders. The ICN contains two kinds of features: spatial patterns and temporal patterns. The spatial maps of ICNs were reliable across resting state sessions based on ICA and dual regression<sup>26</sup>. The complexity of temporal patterns for ICNs exhibited moderate-to-high test-retest reliability under different scan conditions<sup>27</sup>. However, the spatial patterns of ICN included tens of thousands of features, which could increase the complexity of the diagnostic models. Notably, the temporal patterns of ICNs included appropriate number of features, which could reflect the network-wise brain dynamics. There are two types of ICN-based temporal patterns: univariate features within ICN and bivariate features between ICNs. Altered intra-ICN amplitude of low frequency fluctuations and inter-ICN functional connectivity were found in children with ADHD<sup>28</sup>. Moreover, the intra-ICN entropy and inter-ICN synchrony might predict the clinical symptoms for ADHD<sup>29</sup>. However, the dynamic inter-ICN functional connectivity for ADHD remains largely unexplored.

Dynamic functional connectivity reflected the time-varying properties of brain dynamics. The ICNs exhibited dynamic functional connectivity in healthy subjects<sup>30</sup>. Previous evidence found that there were several dynamical states of functional connectivity between ICNs<sup>31</sup>. Altered dynamic functional connectivity patterns were found between eyes-open and eyes-closed conditions<sup>32</sup>. Moreover, the dynamic functional connectivity might underlie spontaneous fluctuations in attention<sup>33</sup>. Notably, the dynamic functional connectivity could successfully discriminate patients with ADHD<sup>34</sup>. However, those studies focused on the temporal clusters of dynamic functional connectivity. Current evidence found that the strength of functional connectivity showed significant fluctuations over time<sup>35</sup>. Given concerns about the temporal variations in time-resolved ICNs, the temporal variability of dynamic functional connectivity still remains unclear.

In this paper, we aimed to build diagnostic model for ADHD based on temporal variability of dynamic functional connectivity. We also sought to find the discriminative patterns of dynamic functional connectivity for ADHD. To achieve these goals, a cohort of children with ADHD and a cohort of healthy controls were obtained from the ADHD-200 Consortium. The diagnostic model was based on SVMs and inter-ICN variability. The performances of diagnostic model and the most discriminative patterns were determined by leave-one-out cross-validation (LOOCV) and 10-folds CVs, respectively. The prediction accuracy of different inter-ICN features were also analyzed for comparison additionally.

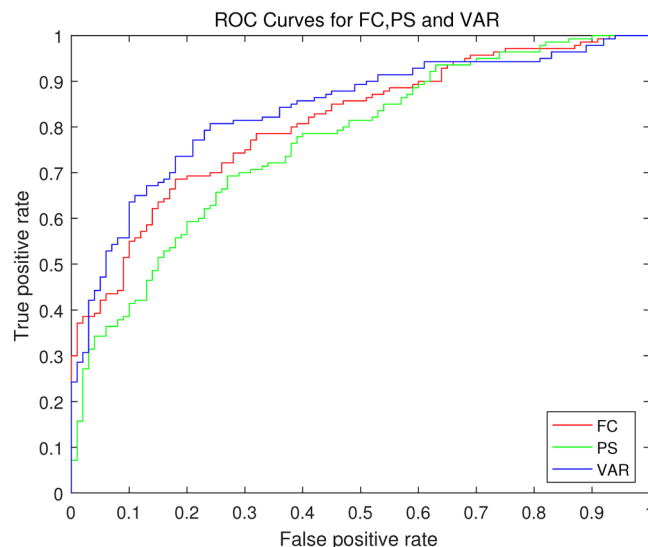
## Results

**Performance of diagnostic models.** The diagnostic model exhibited moderate performance based on LOOCV. Table 1 shows the performance of LOOCV. Figure 1 shows the receiver operating characteristic (ROC) curves based on different measures. The AUC value based on inter-ICN functional connectivity (FC) is 0.81. The area under curve (AUC) value based on inter-ICN phase synchrony (PS) is 0.77. The AUC value based on inter-ICN variability (VAR) is 0.84. Notably, the diagnostic model based on inter-ICN variability outperforms those based on inter-ICN FC or inter-ICN PS.

The diagnostic model exhibited well-established performance based on 10-folds CVs with 1000 times of simulations. Figure 2 shows the histograms for performances of 10-folds classifications. Figure 3 shows the histograms of AUC values based on different measures. The AUC value based on inter-ICN FC is  $0.81 \pm 0.01$ . The AUC value based on inter-ICN PS is  $0.77 \pm 0.01$ . The AUC value based on inter-ICN variability is  $0.81 \pm 0.01$ . Table 2 show the mean performance of 10-folds classifications. Specially, the diagnostic model based on inter-ICN variability exhibits better performance than those based on inter-ICN FC or inter-ICN PS. Table 3 shows the performance of previous methods and our method.

**Discriminative patterns for ADHD based on LOOCV.** Figure 4 shows the ADHD-related discriminative patterns of inter-ICN variability discovered by LOOCV. There are 19 negative feature-weights and 26 positive feature-weights in the ADHD-related diagnostic model. The power weight of inter-ICN variability between RFPN and AN is 1.23, which is the most positive discriminative weight. The power weight of inter-ICN variability between AN and DMN is  $-1.32$ , which is the most negative discriminative weight. Specially, the OVN-related inter-ICN variability exhibits discriminated powers. The power weight of inter-ICN variability between RFPN and MVN is  $-0.99$ . The power weight of inter-ICN variability between OVN and LVN is  $-0.95$ . The power weight of inter-ICN variability between OVN and CBN is  $-0.99$ . The power weight of inter-ICN variability between OVN and AN is  $-0.9$ . The LVN also shows discriminative powers. The power weight of inter-ICN variability between LVN and DMN is  $-0.87$ . The power weight of inter-ICN variability between LVN and CBN is  $-0.88$ . In addition, the power weight of inter-ICN variability between SMN and CBN is 1.22.

**Discriminative patterns for ADHD based on 10-folds CV.** Figure 5 shows the frequency of ADHD-related discriminative patterns of inter-ICN variability discovered by 10-folds CVs. The results of frequencies are the top



**Figure 1.** ROC curves based on LOOCV. The red curve denotes ROC based on inter-ICN functional connectivity (FC). The green curve denotes ROC based on inter-ICN phase synchrony (PS). The blue curve denotes ROC based on inter-ICN variability (VAR).

10 most discriminative features based on 10-folds CVs using 1000 times of simulations. The inter-ICN variability between RFPN and AN appears 946 times. The inter-ICN variability between SMN and CBN appears 838 times. The inter-ICN variability between DMN and AN appears 835 times. The inter-ICN variability between MVN and RFPN appears 599 times. Specially, the OVN exhibits discriminative power during 1000 simulations of 10-folds CVs. The inter-ICN variability between OVN and LVN appears 482 times. The inter-ICN variability between OVN and CBN appears 475 times. The inter-ICN variability between OVN and AN appears 476 times.

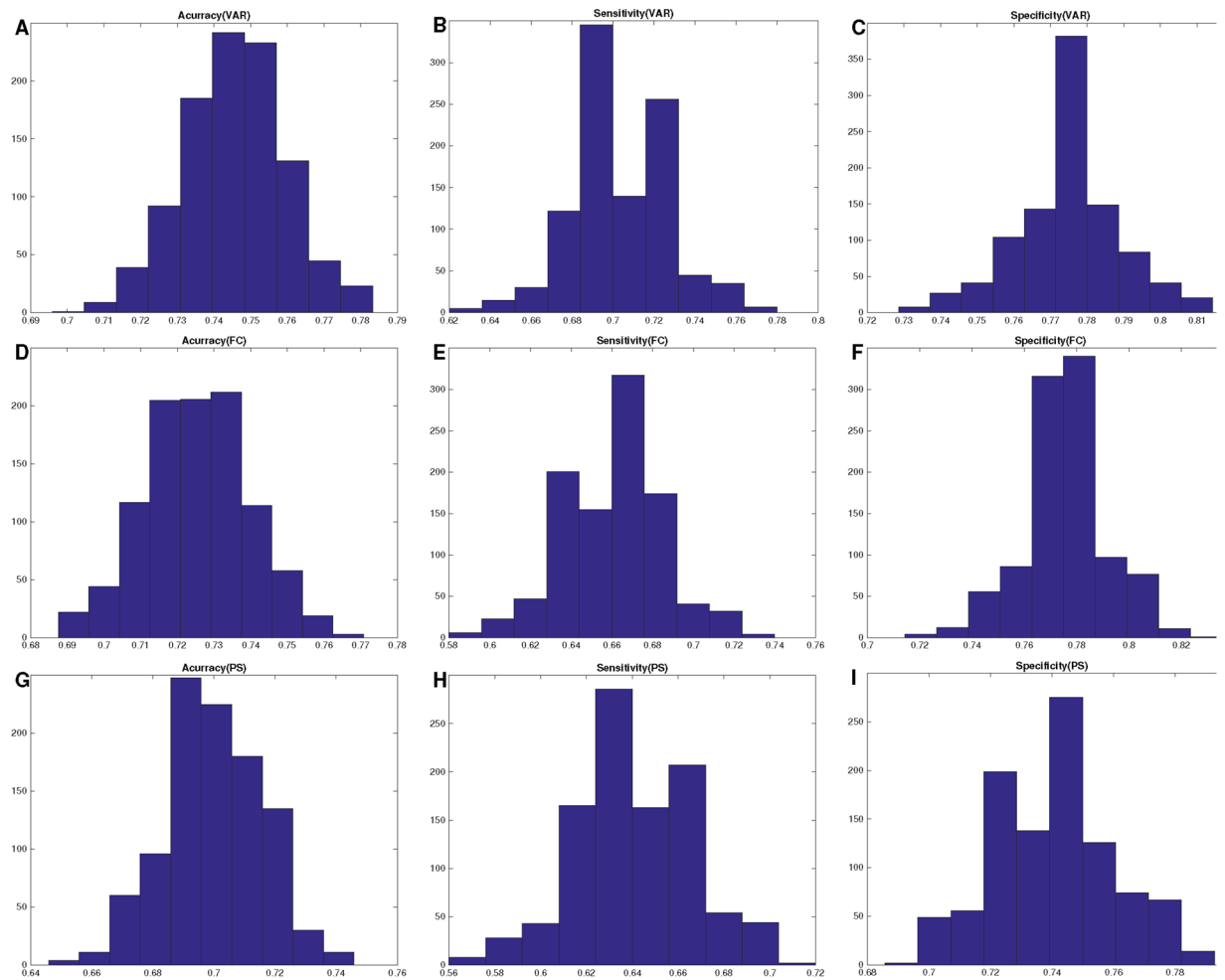
**Discriminative weights of demographic and covariate variables.** Table 4 shows the discriminative weights of demographic and covariate variables during LOOCV and their frequencies during 1000 simulations of 10-folds CVs. The LOOCV reveals that the discriminative weight for gender is 1.63. The discriminative weight for age is 0.35. Notably, the intelligence quotient (IQ) parameters exhibit discriminative power. The discriminative weight for verbal IQ is  $-0.73$ . The discriminative weight for performance IQ is  $-2.17$ . The discriminative weight for full IQ is  $-1.77$ . In addition, the discriminative weight for head motion is 1.63.

Moreover, the diagnostic model is simulated 1000 times using 10-folds CVs. Considering the top 10 most discriminative features, the gender parameter appears 862 times. The verbal IQ parameter appears 143 times. The performance IQ parameter appears 1000 times. The full IQ parameter appears 1000 times. In addition, the head motion parameter appears 994 times during 1000 simulations of 10-folds CV.

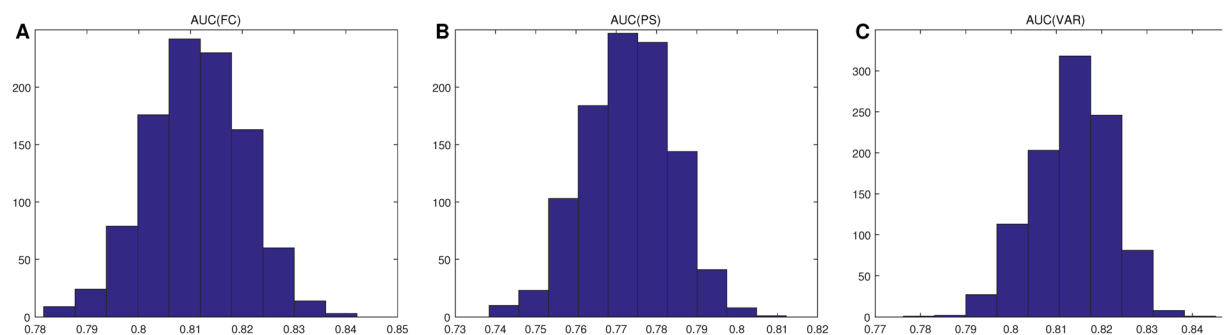
## Discussions

This paper proposed novel inter-ICN variability to identify individuals with ADHD. To the best of our knowledge, this is the first study that investigated the inter-ICN variability for ADHD. The inter-ICN variability was based on the variance of phase differences. The performance of diagnostic model was based on LOOCV and 10-folds CVs. Specially, the 10-folds CVs was simulated 1000 times to reduce overfitting problems. The well-validated results suggested that the diagnostic model could identify patients with ADHD from normal controls. The discriminative patterns of inter-ICN variability was found using LOOCV and 10-folds CVs. In addition, the IQ and head motion parameters were related to ADHD. Together, ADHD could be discriminated from healthy subjects through the inter-ICN variability as well as demographic and covariate variables. The discriminative patterns of inter-ICN variability could be potential biomarkers for ADHD.

**Performance of diagnostic model.** A number of diagnostic models have been built for ADHD based on resting state fMRI. Using a relatively small number of subjects, patients with ADHD could be discriminated from healthy controls based on feature extraction and support vector machines (SVMs)<sup>8,9</sup>. ADHD-200 Consortium provided a significant large number of samples of ADHD and normal controls<sup>10</sup>. Based on ADHD-200 samples, different kinds of features and machine learning methods were applied to diagnosis ADHD<sup>11–13</sup>. Using ADHD-200 samples, the best performance of classification for ADHD is 69.59% based on voxel-wise network features<sup>17</sup>. Total accuracy of 55% was achieved using multimodal features<sup>12</sup>. The prediction accuracy is 67% for combined type and inattentive type of ADHD classification based on combination of functional features<sup>36</sup>. However, most of the current diagnostic models for ADHD exhibited poor performance with either low sensitivity or low specificity. One possible explanation of low prediction accuracy was that the dimensions of neuroimaging features were relatively high in most diagnostic models. The regional cortical surface-based morphological patterns could discriminate ADHD patients from healthy controls based on machine learning methods<sup>14,15</sup>. However, surface-based or voxel-based morphological patterns were structural measures, which could not reflect the brain function or



**Figure 2.** Histograms for performances of 10-folds classifications. Subfigure (A) denotes the total accuracy based on inter-ICN variability (VAR). Subfigure (B) denotes the sensitivity based on inter-ICN variability. Subfigure (C) denotes the specificity based on inter-ICN variability. Subfigure (D) denotes the total accuracy based on inter-ICN functional connectivity (FC). Subfigure (E) denotes the sensitivity based on inter-ICN FC. Subfigure (F) denotes the specificity based on inter-ICN FC. Subfigure (G) denotes the total accuracy based on inter-ICN phase synchrony (PS). Subfigure (H) denotes the sensitivity based on inter-ICN PS. Subfigure (I) denotes the specificity based on inter-ICN PS.



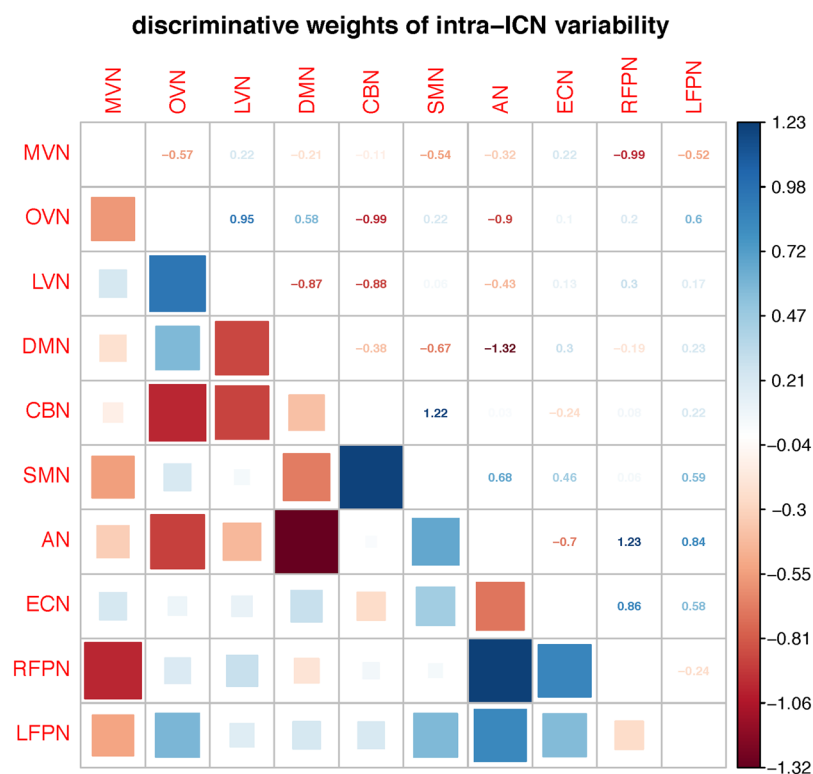
**Figure 3.** Histograms of AUC values based on 10-folds CV. Subfigure (A) denotes histogram of AUC values based on inter-ICN functional connectivity (FC). Subfigure (B) denotes histogram of AUC values based on inter-ICN phase synchrony (PS). Subfigure (C) denotes histogram of AUC values based on inter-ICN variability (VAR).

performance	accuracy	sensitivity	specificity
VAR	75.54% ± 1.34%	70.5% ± 2.34%	77.44% ± 1.47%
FC	72.75% ± 1.43%	65.95% ± 2.46%	77.6% ± 1.59%
PS	70.04% ± 1.58%	64.21% ± 2.58%	74.2% ± 1.88%

**Table 2.** Performance of the diagnostic model based on 10-folds CV.

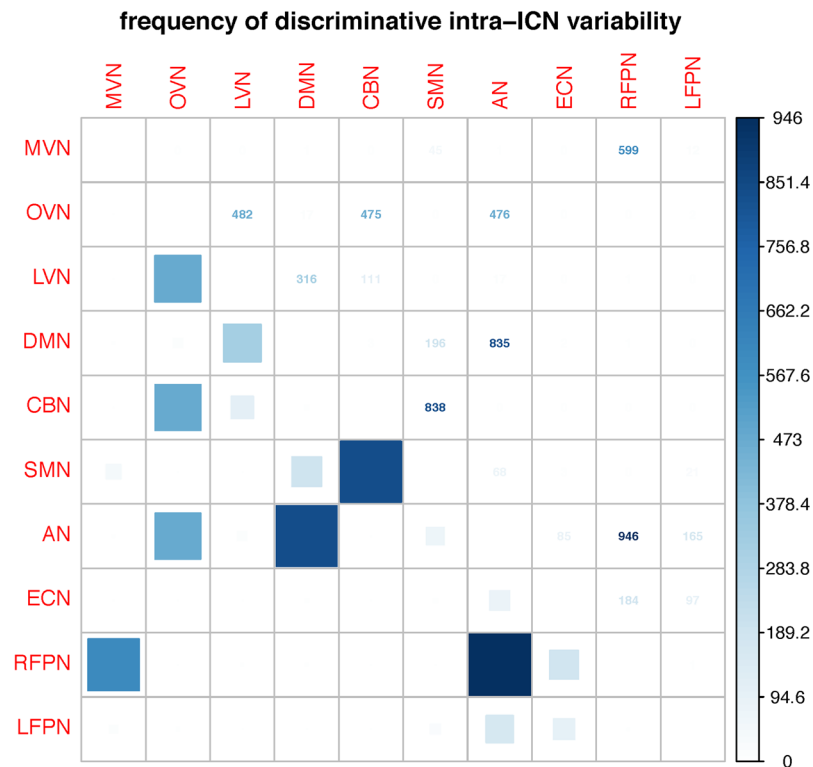
Study	Features	Classifier	Number of features	Subjects (ADHD/TD)	Cross-validation	Acc.	Sen.	Spec.
Zhu <i>et al.</i> <sup>8</sup>	Regional Homogeneity	PCA-FDA	>50k	9/11	LOOCV	85%	78%	91%
Wang <i>et al.</i> <sup>9</sup>	Regional Homogeneity	SVM	>50k	23/23	LOOCV	80%	87%	74%
Dai <i>et al.</i> <sup>11</sup>	Regional Homogeneity	SVM	>50k	222 /402	10 folds	66%	23%	90%
Colby <i>et al.</i> <sup>12</sup>	Multi-modal	SVM	>50k	285/491	10 folds	55%	33%	80%
Cheng <i>et al.</i> <sup>13</sup>	Multi-modal	SVM	>50k	101/143	LOOCV	76%	63%	85%
Peng <i>et al.</i> <sup>14</sup>	Cortical features	ELM	340	55/55	LOOCV	90%	—	—
Qureshi <i>et al.</i> <sup>15</sup>	Cortical features	H-ELM	320	106 /53	10 folds	61%	—	—
Jie <i>et al.</i> <sup>18</sup>	Hyper-connectivity	SVM	—	118/98	LOOCV	83%	84%	82%
Present study	Dynamic FC	SVM	50	100/140	LOOCV	79%	76%	81%
Present study	Dynamic FC	SVM	50	100/140	1000 times of 10 folds	76%	71%	77%

**Table 3.** Performance of previous methods and our method.



**Figure 4.** Discriminative patterns based on LOOCV. Red squares denote decreased inter-ICN variability in ADHD. Blue squares denote increased inter-ICN variability in ADHD.

topology of ADHD. Notably, the diagnostic model based on hyper-connectivity could classify ADHD with total accuracy of 82.9%, sensitivity of 83.9% and specificity of 81.6%<sup>18</sup>. However, the performance of classifiers was not validated using sufficient simulations of 10-folds CVs. Moreover, the mathematical implementation of the aforementioned study was more complex than conventional methods. With novel inter-ICN variability, this study could identify patients with ADHD based on LOOCV and 10-folds CVs. Using LOOCV, the total accuracy was 78.75%, the sensitivity was 76%, and specificity was 80.71%. Using 10-folds CVs, the average prediction accuracy was 75.54% ± 1.34%, the average sensitivity was 70.5% ± 2.34%, and the average specificity was 77.44% ± 1.47%.



**Figure 5.** Frequency of discriminative patterns based on 10-folds CVs. The numbers and squares denote the frequency of top 10 discriminative patterns based on 10-folds CV.

	LOOCV	10-folds CV
Gender	1.63	1000
Age	0.35	0
Verbal IQ	-0.73	143
Performance IQ	-2.17	1000
Full IQ	-1.77	1000
FD	1.63	994

**Table 4.** Discriminative powers of demographic variables.

The diagnostic model exhibited balanced accuracies of sensitivity and specificity. Therefore, the diagnostic model could be beneficial for clinical applications. In addition, we proposed a novel neural metrics with relatively low dimensions of features, which could reduce the complexity of the diagnostic models. Specially, the results were based on 1000 times 10-folds CV, while most of the previous methods were based on LOOCV or one run of 10-folds CV. In summary, our method could open a new perspective for diagnosing ADHD.

**Discriminative patterns of dynamic functional connectivity for ADHD.** Discriminative patterns of inter-ICN variability were found for ADHD. In the discriminative model, the positive weights indicated the increased inter-ICN variability in ADHD, while the negative weights indicated the decreased inter-ICN variability in ADHD. Most of the discriminative patterns were with positive weights, which suggested increased inter-network activity in ADHD. Previous evidence found that there were enhanced brain activities in ADHD than normal controls<sup>4</sup>. Altered brain topologies were found in ADHD based on graphical measures<sup>37</sup>. Of note, our previous study also found altered brain topologies based on inter-ICN functional connectivity<sup>28</sup>. In this study, the diagnostic model identified certain features of variability between ICNs (i.e., AN-RFPN, AN-DMN) that exhibited discriminative powers. The DMN-related cortex exhibited altered resting state regional homogeneity<sup>9</sup>. Enhanced DMN-related activities were correlated to sustained attention deficits<sup>38</sup>. The RFPN-related variability might be related to the prefrontal-striatal-cerebellar circuit in ADHD<sup>39</sup>. In addition, the visual-related ICNs (i.e., MVN, OVN) were found with discriminative powers. The ICN-based evidences noted that the sensory- and visual-related brain networks might contribute to ADHD<sup>28</sup>. Previous evidence also found enhanced brain activities in sensory-related brain regions<sup>4</sup>. In summary, the informative discriminative patterns could be beneficial to diagnostic model of ADHD.



**Temporal variability of dynamic functional connectivity.** The temporal variability of dynamic functional connectivity exhibited better discriminative power than conventional measures of connectivity. A possible explanation was that the inter-ICN variability might reflect the dynamic properties of brain networks. The inter-ICN functional connectivity and phase synchrony was static measures, which contained less information of brain dynamics than inter-ICN variability. Previous evidence suggested that the brain should exhibit dynamic activities during resting state. Seven temporal clusters of dynamic functional connectivity were found based on sliding-window and machine learning methods<sup>31</sup>. Moreover, the dynamic functional connectivity between ICNs were related to visual attention (i.e., eyes-open/closed states)<sup>32</sup>. However, the aforementioned studies were based on temporal clustering, which could not reflect the temporal fluctuations of dynamic functional connectivity. A previous study applied amplitude of low frequency fluctuations (ALFF) on temporal dynamic functional connectivity to predict brain maturity<sup>40</sup>. However, the results were based on sliding-window method, which has certain limitations (i.e., manually selection of sliding-window length). Given current concerns on temporal variability of dynamic functional connectivity, the proposed inter-ICN variability was based on the phase differences, which did not depend on sliding-window. Thus, the inter-ICN variability might be potential biomarkers for human connectome.

**IQ and ADHD.** Previous evidences found that the IQ was related to ADHD. The developments of cortical thickness and white matters might be delayed in ADHD with low IQ<sup>41</sup>. The IQ scores also correlated to ALFFs within ICNs, implying that the IQ might contribute to ADHD<sup>28</sup>. A recent study found that functional connectivity between ICNs could predict the IQ scores for ADHD<sup>42</sup>. In this paper, the IQ scores exhibited discriminative power for ADHD, suggesting that the IQ might play an important role in ADHD. Without the IQ scores, the diagnostic model achieved total accuracy of 71.67% based on LOOCV, which was much lower than that with the IQ scores. The personal characteristic features (i.e., IQ) could outperform neuroimaging data in diagnostic models for ADHD<sup>43</sup>. Moreover, the IQ scores were negatively related to ADHD in our diagnostic model. The results implied that the ADHD group exhibited lower IQ scores than healthy group.

**Head motion and ADHD.** Given current concerns on head motion in resting state fMRI, the head motion parameter was taken as supplementary feature in the diagnostic model. The power weight of head motion was 1.63, which means that patients exhibited more head motion than normal controls. Without head motion parameter, the diagnostic model achieved total accuracy of 74.17% using LOOCV, which was lower than that with head motion. Thus, ADHD might possess distinctive head motion during resting state. The head motion has been identified as an important confound in resting state fMRI. Previous studies suggested that the subjects with Frame-Displacement (FD) more than 0.5 mm should be discarded<sup>44,45</sup>, since the functional signals were corrupted with severe head motions. The original sample size for this study was 245 subjects. Only 5 subjects exhibited severe head motion, resulting in 240 subjects for diagnostic model. Thus, the preprocessed subjects could represent the original samples. Besides, we used FD as a covariate in the diagnostic models. Previous studies suggested that head motion should be considered as a covariate variable in statistical models<sup>44,45</sup>. Moreover, head motion might affect the test-retest reliability of resting state features<sup>27,46</sup>. For group-wise analysis, the mean FD was computed as head motion parameter<sup>47,48</sup>. In this study, head motion significantly improved the performance of diagnostic model. The results suggested that head motion should be carefully considered in ADHD research.

**Advantages.** One advantage of this study was applying inter-ICN variability to diagnosis ADHD. The feature dimensions of inter-ICN variability was much lower than regional- or voxel-wise measures. Moreover, the inter-ICN variability outperformed conventional connectivity estimators in the diagnostic model. Thus, the inter-ICN variability could be beneficial for classification of ADHD. Another advantage of this study was applying 1000 simulations of 10-folds CVs on the diagnostic models. The results were well-validated compared to previous studies. Moreover, the prediction accuracy was balanced with equal sensitivity and specificity. The third advantage of this study was applying SVMs to discover the discriminative patterns for ADHD. The linear SVMs found certain discriminative features, which could be potential biomarkers for ADHD. The fourth advantage of this study was applying the IQ scores as the predictors in the diagnostic models. The IQ scores were related to ADHD, and could significantly improve the performances of the diagnostic models. The fifth advantage of this study was applying head motion as a predictor in the diagnostic models. The head motion parameter was found to be associated with ADHD. The diagnostic models exhibited better performances with head motion parameter. Overall, the inter-ICN variability, IQ scores and head motion might contribute to diagnosis of ADHD. Moreover, the diagnostic models with balanced performances was well-validated using 1000 simulations of 10-folds CVs.

**Limitations.** There were several limitations which should be noted in this study. One limitation was that the neuroimaging datasets were obtained using different scan parameters. There were more than three different scan protocols for this dataset. Due to the limited information for this public datasets, we introduced the scan difference (i.e., Peking\_1, Peking\_2, Peking\_3, Peking\_1\_test) as a covariate. However, the performance of diagnostic model with scan differences was a little lower than that without scan differences. The diagnostic model with scan differences exhibited accuracy of 76.67%, sensitivity of 72%, and specificity of 80%. One possible explanation was that there were five different scan protocols for anatomical images in Peking\_3. Thus, the covariate of scan difference might be mislabeled in the diagnostic models. The effects of scan parameters should be taken into account in future study. Another limitation was that the diagnostic model was based on Chinese children in order to remove the effects of populations and versions of clinical scales. We tested our diagnostic model on datasets from the NYU site. However, the diagnostic model only achieved total accuracy of 54.08%. We also applied our method on the NYU children using LOOCV, which achieved total accuracy of 58.67%. One explanation was the training and testing models were based on different populations. Another explanation was the different scan

	ADHD	Normal	<i>p</i> -value
Number of subjects	100	140	—
Gender (male: female)	88:12	81:59	<10 <sup>-3</sup>
Handless (R: L)	140:0	140:0	1
Age (year)	12.1 ± 2.05	11.44 ± 1.86	0.0093
Verbal IQ	111.44 ± 15.57	120.46 ± 13.3	<10 <sup>-5</sup>
Performance IQ	99.03 ± 13.65	111.29 ± 14.33	<10 <sup>-9</sup>
Full IQ	106.36 ± 13.02	118.02 ± 12.12	<10 <sup>-10</sup>

**Table 5.** Subjects' demographic and covariate variables.

parameters of the two datasets. The third explanation was the two datasets were based on different versions of clinical scales. Therefore, we should test our model on an independent dataset of Chinese children in future study. The parameters of classifiers should be optimized to improve the performance of the diagnostic model. The third limitation was not considering the subtypes in the diagnostic model, since the subtypes contained the label information. The diagnostic model for combined type of ADHD and normal controls exhibited total accuracy of 69.44%, sensitivity of 73.0% and specificity of 65.7%. The diagnostic model for inattentive type of ADHD and normal controls exhibited total accuracy of 77.3%, sensitivity of 77.4% and specificity of 77.1%. The fourth limitation was based on single modal of neuroimaging features. The performance of the diagnostic model could be improved with multi-modal imaging methods (i.e., anatomical MRI, diffusion MRI, arterial spin labeling MRI). The fifth limitation was not considering the eyes-open/closed conditions in the diagnostic model. The eyes-open/closed conditions could affect functional connectivity. The subjects in this study were asked to have their eyes opened or closed. However, we cannot control this confounding variable due to limited information provided by ADHD-200 website. Additionally, the inter-ICN biomarkers should be validated using an independent dataset for clinical approach. Thus, we plan to test the diagnostic model using different features and populations in further study. We also plan to obtain multimodal imaging datasets in further study. Moreover, the performance of the diagnostic model should be improved using different machine learning methods (i.e., extreme learning machines, Bayesian-based classifiers, and neural networks).

## Conclusion

This paper investigated the temporal variability of dynamic functional connectivity to diagnosis children with ADHD based on machine learning methods. The diagnostic models could discriminate patients with ADHD using cross-validations. Moreover, the discriminative patterns of inter-ICN variability were discovered by the diagnostic model. In summary, individuals with ADHD could be identified by machine learning based on the inter-ICN variability, which could be potential biomarkers for ADHD.

## Methods

**Participants and MRI protocols.** The ADHD-200 Consortium provided a great number of individuals with ADHD and normal controls ([http://fcon\\_1000.projects.nitrc.org/indi/adhd200/](http://fcon_1000.projects.nitrc.org/indi/adhd200/)). ADHD-200 Consortium were consisted of datasets from eight sites (i.e., New York University, Peking University, etc.). To remove the effects of sites, the participants in this study were based on datasets from the Peking site. There were 245 participants using datasets combined from Peking\_1, Peking\_2, Peking\_3 and Peking\_1\_test. The demographic and covariate information could be found in Table 5. The participants were diagnosed using ADHD Rating Scale (ADHD-RS) IV. Participants were selected using the following criteria: 1) right-handedness; 2) without loss of consciousness caused by head trauma; 3) without history of neurological disease and other mental disorder (i.e., schizophrenia, affective disorder, pervasive development disorder, or substance abuse). In addition, the IQ score for each subject was greater than 80 evaluated by Wechsler Intelligence Scale for Chinese Children-Revised (WISCC-R). All research of this study was approved by the Research Ethics Review Board of the Institute of Mental Health, Peking University. The guidelines and regulations of experiments were carried out in accordance with institutional review boards of the Institute of Mental Health, Peking University. Informed consent was provided by the parent of each participant and all of the children agreed to participate in this research.

Both anatomical MRI and resting state fMRI were scanned for each subject. The neuroimaging datasets were obtained from a SIEMENS TRIOTIM syngo 3-T MRI scanner. The anatomical MRI data was a high-resolution T1-weighted MPRAGE 3D volume, which was defaced to protect patient identity. The parameters of the anatomical MRI could be found in Table 6. The resting state fMRI data was a standard T2-weighted EPI 4D volume. The parameters of the functional MRI could be found in Table 7. During each scan, the participants were required to keep relax, stay still with their eyes either open or closed. A black screen with a white fixation cross was presented to each participant during the scan. The additional information could be obtained from the website of ADHD-200 Consortium.

**Data preprocessing.** The anatomical and functional MRI datasets were preprocessed using scripts from 1000 functional connectome project. The preprocessing scripts were based on AFNI and FSL. The anatomical MRI datasets were skull-stripped, segmented into three kinds of brain tissues (i.e., white matter, gray matter, and cerebrospinal fluid), nonlinearly deformed into standard MNI brain space. After discarding the first five volume, the resting state fMRI datasets were preprocessed using the following steps: slice-timing correction, motion



	Peking_1	Peking_2	Peking_3	Peking_1_test
TR/TE (ms)	2530/3.39	2530/3.45	5 protocols	2530/3.39
Slices	128	176	5 protocols	128
Thickness(mm)	1.33	1	5 protocols	1.33
FoV read(mm)	256	256	5 protocols	256
Fov phase	100%	81.3%	5 protocols	100%
Flip angel (degree)	7	7	5 protocols	7

**Table 6.** Scan parameters for anatomical MRIs.

	Peking_1	Peking_2	Peking_3	Peking_1_test
TR/TE (ms)	2000/30	2000/30	2000/30	2000/30
Slices	33	33	30	33
Thickness(mm)	3.5	3	4.5	3.5
FoV read(mm)	200	200	220	200
Fov phase	100%	100%	—	100%
Flip angel (degree)	90	90	90	90
volumes	240	240	240	239

**Table 7.** Scan parameters for functional MRIs.

index	Names of ICNs	Abbreviations
ICN1	Medial visual network	MVN
ICN2	Occipital visual network	OVN
ICN3	Lateral visual network	LVN
ICN4	Default mode network	DMN
ICN5	Cerebellum network	CBN
ICN6	Sensorimotor network	SMN
ICN7	Auditory network	AN
ICN8	Executive control network	ECN
ICN9	Right frontoparietal network	RFPN
ICN10	Left frontoparietal network	LFPN

**Table 8.** Names of template ICNs.

correction, skull stripping, regressing out nuisance signals (i.e., white matter, cerebrospinal fluid, and Friston-24 motion parameters) as well as linear and quantic trends, nonlinearly wrapped into standard brain space, band-pass filtering (0.01–0.1 Hz), spatially smoothing (FWHM = 6 mm). Of note, after motion correction, the mean frame-wise displacement (FD) was computed as the motion covariate for each subject. Five participants with mean FD larger than 0.5 mm were discarded, resulting in 240 subjects for diagnostic model. According to statistics, 37 subjects (37%) were combined type of ADHD. 1 subject (1%) was hyperactive/impulsive type of ADHD. 62 subjects (62%) were inattentive type of ADHD.

**Time-courses of ICNs.** In this study, ten well-established ICNs were taken as spatial templates. The template ICNs were obtained using meta-analysis of the BrainMap database. The names and abbreviations of the ten ICNs were present in Table 8. The time-course of each ICN was computed using the spatial regression step of dual-regression. First, the 3D image of the corresponding ICN was extracted from the ten template ICNs. Second, the 3D image of a functional volume was extracted from the 4D resting state fMRI sequence. Third, the 3D images of ICN and functional volume were reshaped into two one-dimensional vectors. Fourth, the beta value was obtained by spatial regression between the two vectors. Finally, the time-courses of each ICN were obtained by combining the beta values along the timeline of the functional sequence. The detailed steps of spatial regression could be found in previous studies<sup>22,49</sup>.

**Inter-ICN variability.** The inter-ICN variability was based on Hilbert transform with the following procedures: 1) obtain pair-wise time-courses of ICNs; 2) apply Hilbert transform on the two time-courses; 3) obtain the instantaneous phases of each time-course; 4) compute the instantaneous phase differences between the two time-courses; 5) transform instantaneous phase differences into  $-\pi$  to  $\pi$ ; 6) the compute the variance of the instantaneous phase differences. Finally, after looping through the ten ICNs, a vector with  $10 \times 9/2 = 45$  features were obtained for each subject. In addition, the conventional functional connectivity (FC) and phase synchrony (PS) were computed for comparisons.

**Diagnostic model.** The original features were consisted of inter-ICN variability, IQ scores, age, sex and head motion. The diagnostic model could be denoted as the following formula:

$$label = \sum_{i=1}^{45} w_{1i} VAR_i + \sum_{i=1}^3 w_{2i} IQ_i + w_3 sex + w_4 age + w_5 mFD + b$$

In the diagnostic model, the label denotes the diagnostic information of each subject (i.e., 1 for ADHD, 0 for normal controls). The  $VAR_i$  denotes the  $i$ th inter-ICN variability, and  $w_{1i}$  denotes the weight of the  $i$ th inter-ICN variability. The  $IQ_i$  denotes verbal IQ, performance IQ or full IQ, and  $w_{2i}$  denotes the weight the  $i$ th  $IQ_i$ . The sex denotes the gender of subject, and  $w_3$  denotes the weight of sex. The age denoted the biological age of subject, and  $w_4$  denotes the weight of age. The mFD denotes mean frame-wise displacement, and  $w_5$  denotes the weight of head motion.

The diagnostic model was analyzed using support vector machines (SVMs), which was proposed by Cortes and Vapnik in 1995<sup>50</sup>. SVM was designed for classification of two classes. The advantage of SVM was solving small sample problem, nonlinear problem and high dimensional pattern recognition. The basic idea of SVM was searching an optimized hyperplane which can classify different kinds of samples. In this paper, sequential minimal optimization (SMO) was applied on training and testing datasets to search the optimized hyperplane. SMO was an iterative algorithm with fast speed to effectively solve the optimizing problems of SVMs<sup>51</sup>. The trick of SMO was analytically solving a set of smallest possible sub-problems instead of the original SVMs. SMO significantly improve the performance and computation times of SVMs. Moreover, SMO exhibits good performance for linear SVMs. Here, SMO with linear kernels was applied in the diagnostic model. The implementation of SMO procedure was based on WEKA, which is a popular machine learning software ([www.cs.waikato.ac.nz/ml/weka](http://www.cs.waikato.ac.nz/ml/weka)). In addition, the weights of features were considered as their contributions to the diagnostic model<sup>52</sup>. The positive weight means increased inter-ICN variability in ADHD, while the negative weight means decreased inter-ICN variability in ADHD.

**Performance of diagnostic model.** To evaluate the performance of diagnostic model, cross-validation (CV) was applied in this study. In an  $n$ -fold CV, the original samples were first divided into  $n$ -folds. Then, one-fold of samples were selected as testing samples, leaving the rest instances as training samples. Third, a diagnostic model was built on the training samples. Fourth, the trained diagnostic model was tested using testing samples. Finally,  $n$  diagnostic models were trained and tested based on  $n$ -fold CV. In this paper, leave-one-out cross-validation (LOOCV) and 10-folds CV were applied to evaluate the performance of diagnostic model. Specially, 1000 simulations of 10-fold CV were performed to validate the effects of random seeds in partitions of folds.

The performance of classification was evaluated by total accuracy (Acc.), sensitivity (Sen.) and specificity (Spec.). Here, let TP denotes the number of children with ADHD correctly classified as patients. FP denotes the number of healthy controls incorrectly classified as ADHD. TN denotes the number of correctly identified healthy subjects. FN denotes the number of incorrectly identified patients. Total accuracy is defined as the proportion of correctly predicted instances (i.e.,  $accuracy = (TN + TP)/(TN + FN + TP + FP)$ ). Sensitivity is the proportion of correctly classified positive instances (i.e.,  $sensitivity = TP/(TP + FN)$ ). Specificity is the proportion of correctly classified negative instances (i.e.,  $specificity = TN/(TN + FP)$ ).

**Data availability.** The MRI datasets could be obtained from a public database ([http://fcon\\_1000.projects.nitrc.org/indi/adhd200/](http://fcon_1000.projects.nitrc.org/indi/adhd200/)).

## References

1. APA. *Diagnostic and statistical manual of mental disorders: DSM-IV-TR*, 943 (Springer, 2000).
2. Biederman, J. Attention-deficit/hyperactivity disorder: a selective overview. *Biological psychiatry* 57, 1215–1220, <https://doi.org/10.1016/j.biopsych.2004.10.020> (2005).
3. Kelly, A. M. C., Margulies, D. S. & Castellanos, F. X. Recent advances in structural and functional brain imaging studies of attention-deficit/hyperactivity disorder. *Curr Psychiatry Rep* 9, 401–407 (2007).
4. Tian, L. *et al.* Enhanced resting-state brain activities in ADHD patients: a fMRI study. *Brain Dev* 30, 342–348, <https://doi.org/10.1016/j.braindev.2007.10.005> (2008).
5. Uddin, L. Q. *et al.* Network homogeneity reveals decreased integrity of default-mode network in ADHD. *J Neurosci Methods* 169, 249–254, <https://doi.org/10.1016/j.jneumeth.2007.11.031> (2008).
6. Cao, X. *et al.* Abnormal resting-state functional connectivity patterns of the putamen in medication-naïve children with attention deficit hyperactivity disorder. *Brain Res* 1303, 195–206, <https://doi.org/10.1016/j.brainres.2009.08.029> (2009).
7. Qiu, A. *et al.* Basal ganglia volume and shape in children with attention deficit hyperactivity disorder. *The American journal of psychiatry* 166, 74–82, <https://doi.org/10.1176/appi.ajp.2008.08030426> (2009).
8. Zhu, C.-Z. *et al.* Fisher discriminative analysis of resting-state brain function for attention-deficit/hyperactivity disorder. *NeuroImage* 40, 110–120, <https://doi.org/10.1016/j.neuroimage.2007.11.029> (2008).
9. Wang, X., Jiao, Y., Tang, T., Wang, H. & Lu, Z. Altered regional homogeneity patterns in adults with attention-deficit hyperactivity disorder. *Eur J Radiol* 82, 1552–1557, <https://doi.org/10.1016/j.ejrad.2013.04.009> (2013).
10. Milham, M. P., Fair, D., Mennes, M. & Mostofsky, S. H. The ADHD-200 Consortium: A Model to Advance the Translational Potential of Neuroimaging in Clinical Neuroscience. *Front Syst Neurosci* 6, 1–5, <https://doi.org/10.3389/fnsys.2012.00062> (2012).
11. Dai, D., Wang, J., Hua, J. & He, H. Classification of ADHD children through multimodal magnetic resonance imaging. *Front Syst Neurosci* 6, 1–8, <https://doi.org/10.3389/fnsys.2012.00063> (2012).
12. Colby, J. B. *et al.* Insights into multimodal imaging classification of ADHD. *Front Syst Neurosci* 6, 59, <https://doi.org/10.3389/fnsys.2012.00059> (2012).
13. Cheng, W., Ji, X., Zhang, J. & Feng, J. Individual classification of ADHD patients by integrating multiscale neuroimaging markers and advanced pattern recognition techniques. *Front Syst Neurosci* 6, 58, <https://doi.org/10.3389/fnsys.2012.00058> (2012).
14. Peng, X., Lin, P., Zhang, T. & Wang, J. Extreme learning machine-based classification of ADHD using brain structural MRI data. *PLoS one* 8, e79476, <https://doi.org/10.1371/journal.pone.0079476> (2013).

15. Qureshi, M. N., Min, B., Jo, H. J. & Lee, B. Multiclass Classification for the Differential Diagnosis on the ADHD Subtypes Using Recursive Feature Elimination and Hierarchical Extreme Learning Machine: Structural MRI Study. *PLoS one* **11**, e0160697, <https://doi.org/10.1371/journal.pone.0160697> (2016).
16. Dey, S., Rao, A. R. & Shah, M. Attributed graph distance measure for automatic detection of attention deficit hyperactive disordered subjects. *Front Neural Circuits* **8**, 64, <https://doi.org/10.3389/fncir.2014.00064> (2014).
17. Dey, S., Rao, A. R. & Shah, M. Exploiting the brain's network structure in identifying ADHD subjects. *Front Syst Neurosci* **6**, 75, <https://doi.org/10.3389/fnsys.2012.00075> (2012).
18. Jie, B., Wee, C. Y., Shen, D. & Zhang, D. Hyper-connectivity of functional networks for brain disease diagnosis. *Med Image Anal* **32**, 84–100, <https://doi.org/10.1016/j.media.2016.03.003> (2016).
19. Fox, M. D. *et al.* The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proc Natl Acad Sci USA* **102**, 9673–9678, <https://doi.org/10.1073/pnas.0504136102> (2005).
20. Fox, M. D. & Raichle, M. E. Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Nat Rev Neurosci* **8**, 700–711, <https://doi.org/10.1038/nrn2201> (2007).
21. Beckmann, C. F., DeLuca, M., Devlin, J. T. & Smith, S. M. Investigations into resting-state connectivity using independent component analysis. *Philos Trans R Soc Lond B Biol Sci* **360**, 1001–1013, <https://doi.org/10.1098/rstb.2005.1634> (2005).
22. Smith, S. M. *et al.* Correspondence of the brain's functional architecture during activation and rest. *Proc Natl Acad Sci USA* **106**, 13040–13045, <https://doi.org/10.1073/pnas.0905267106> (2009).
23. Biswal, B. B. *et al.* Toward discovery science of human brain function. *Proc Natl Acad Sci USA* **107**, 4734–4739, <https://doi.org/10.1073/pnas.0911855107> (2010).
24. Laird, A. R. *et al.* Behavioral interpretations of intrinsic connectivity networks. *J Cogn Neurosci* **23**, 4022–4037 (2011).
25. Damoiseaux, J. S. *et al.* Consistent resting-state networks across healthy subjects. *Proc Natl Acad Sci USA* **103**, 13848–13853, <https://doi.org/10.1073/pnas.0601417103> (2006).
26. Zuo, X.-N. *et al.* Reliable intrinsic connectivity networks: test-retest evaluation using ICA and dual regression approach. *NeuroImage* **49**, 2163–2177, <https://doi.org/10.1016/j.neuroimage.2009.10.080> (2010).
27. Wang, X., Jiao, Y., Tang, T., Wang, H. & Lu, Z. Investigating univariate temporal patterns for intrinsic connectivity networks based on complexity and low-frequency oscillation: a test-retest reliability study. *Neuroscience* **254**, 404–426, <https://doi.org/10.1016/j.neuroscience.2013.09.009> (2013).
28. Wang, X. H. & Li, L. Altered temporal features of intrinsic connectivity networks in boys with combined type of attention deficit hyperactivity disorder. *Eur J Radiol* **84**, 947–954, <https://doi.org/10.1016/j.ejrad.2015.02.018> (2015).
29. Wang, X.-H., Jiao, Y. & Li, L. Predicting clinical symptoms of attention deficit hyperactivity disorder based on temporal patterns between and within intrinsic connectivity networks. *Neuroscience* **362**, 60–69, <https://doi.org/10.1016/j.neuroscience.2017.08.038> (2017).
30. Hutchison, R. M., Womelsdorf, T., Gati, J. S., Everling, S. & Menon, R. S. Resting-state networks show dynamic functional connectivity in awake humans and anesthetized macaques. *Hum Brain Mapp* **2177**, 2154–2177, <https://doi.org/10.1002/hbm.22058> (2012).
31. Allen, E. A. *et al.* Tracking Whole-Brain Connectivity Dynamics in the Resting State. *Cerebral cortex* **24**, 663–676, <https://doi.org/10.1093/cercor/bhs352> (2014).
32. Wang, X. H., Li, L., Xu, T. & Ding, Z. Investigating the Temporal Patterns within and between Intrinsic Connectivity Networks under Eyes-Open and Eyes-Closed Resting States: A Dynamical Functional Connectivity Study Based on Phase Synchronization. *PLoS one* **10**, e0140300, <https://doi.org/10.1371/journal.pone.0140300> (2015).
33. Kucyi, A., Hove, M. J., Esterman, M., Hutchison, R. M. & Valera, E. M. Dynamic Brain Network Correlates of Spontaneous Fluctuations in Attention. *Cerebral cortex*, <https://doi.org/10.1093/cercor/bhw029> (2016).
34. Ou, J. *et al.* Atomic dynamic functional interaction patterns for characterization of ADHD. *Human brain mapping* **35**, 5262–5278, <https://doi.org/10.1002/hbm.22548> (2014).
35. Zalesky, A., Fornito, A., Cocchi, L., Gollo, L. L. & Breakspear, M. Time-resolved resting-state brain networks. *Proc Natl Acad Sci USA* **111**, 10341–10346, <https://doi.org/10.1073/pnas.1400181111> (2014).
36. Sato, J. J. R., Hoexter, M. Q. M., Fujita, A. & Rohde, L. A. Evaluation of pattern recognition and feature extraction methods in ADHD prediction. *Front Syst Neurosci* **6**, 1–14, <https://doi.org/10.3389/fnsys.2012.00068> (2012).
37. Wang, L. *et al.* Altered small-world brain functional networks in children with attention-deficit/hyperactivity disorder. *Hum Brain Mapp* **30**, 638–649, <https://doi.org/10.1002/hbm.20530> (2009).
38. Bonnelle, V. *et al.* Default mode network connectivity predicts sustained attention deficits after traumatic brain injury. *J Neurosci* **31**, 13442–13451, <https://doi.org/10.1523/JNEUROSCI.1163-11.2011> (2011).
39. Castellanos, F. X. & Proal, E. Large-scale brain systems in ADHD: beyond the prefrontal-striatal model. *Trends Cogn Sci* **16**, 17–26, <https://doi.org/10.1016/j.tics.2011.11.007> (2012).
40. Qin, J. *et al.* Predicting individual brain maturity using dynamic functional connectivity. *Frontiers in Human Neuroscience* **9**, <https://doi.org/10.3389/fnhum.2015.00418> (2015).
41. de Zeeuw, P. *et al.* Differential brain development with low and high IQ in attention-deficit/hyperactivity disorder. *PLoS one* **7**, e35770, <https://doi.org/10.1371/journal.pone.0035770> (2012).
42. Park, B.-y., Hong, J., Lee, S.-H. & Park, H. Functional Connectivity of Child and Adolescent Attention Deficit Hyperactivity Disorder Patients: Correlation with IQ. *Frontiers in Human Neuroscience* **10**, <https://doi.org/10.3389/fnhum.2016.00565> (2016).
43. Brown, M. R. *et al.* ADHD-200 Global Competition: diagnosing ADHD using personal characteristic data can outperform resting state fMRI measurements. *Front Syst Neurosci* **6**, 69, <https://doi.org/10.3389/fnsys.2012.00069> (2012).
44. Yan, C.-G. *et al.* A comprehensive assessment of regional variation in the impact of head micromovements on functional connectomics. *NeuroImage* **76**, 183–201, <https://doi.org/10.1016/j.neuroimage.2013.03.004> (2013).
45. Yan, C.-G., Craddock, R. C., He, Y. & Milham, M. P. Addressing head motion dependencies for small-world topologies in functional connectomics. *Front Hum Neurosci* **7**, 910, <https://doi.org/10.3389/fnhum.2013.00910> (2013).
46. Zuo, X.-N. *et al.* Toward reliable characterization of functional homogeneity in the human brain: Preprocessing, scan duration, imaging resolution and computational space. *NeuroImage* **65**, 374–386, <https://doi.org/10.1016/j.neuroimage.2012.10.017> (2013).
47. 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, <https://doi.org/10.1016/j.neuroimage.2011.10.018> (2012).
48. Power, J. D. *et al.* Methods to detect, characterize, and remove motion artifact in resting state fMRI. *NeuroImage* **84**, 320–341, <https://doi.org/10.1016/j.neuroimage.2013.08.048> (2014).
49. Filippini, N. *et al.* Distinct patterns of brain activity in young carriers of the APOE-epsilon4 allele. *Proc Natl Acad Sci USA* **106**, 7209–7214, <https://doi.org/10.1073/pnas.0811879106> (2009).
50. Cortes, C. & Vapnik, V. N. Support-Vector Networks. *Machine Learning* **297**, 273–297 (1995).
51. Platt, J. C. In *Advances in kernel methods* 185–208 (MIT Press, 1999).
52. Ecker, C. *et al.* Investigating the predictive value of whole-brain structural MR scans in autism: a pattern classification approach. *NeuroImage* **49**, 44–56, <https://doi.org/10.1016/j.neuroimage.2009.08.024> (2010).

## Acknowledgements

This work was supported in part by grants from the National Natural Science Foundation of China (61601155, 81501453, 61731008, and 61271063), 973 Program (2013CB329502), the National Distinguished Young Research Scientist Award (60788101), and the Zhejiang Provincial Natural Science Foundation of China (LZ15F010001). We thank the ADHD-200 Consortium for sharing brain imaging datasets of ADHD and normal controls.

## Author Contributions

X.H.W. and L.L. contributed to the idea of the paper. Y.J. performed data preprocessing and statistical analysis. X.H.W. wrote the main manuscript text and prepared figures and tables. All authors contributed to the paper and reviewed the final manuscript.

## Additional Information

**Competing Interests:** The authors declare no competing interests.

**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 license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license 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 license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018