

Supplementary Information: Diffusion trajectory of atypical morphological development in autism spectrum disorder

Yu Feng^{1,2}, Weixing Zhao^{1,2}, Youyi Li^{1,2}, Qin Yin^{1,2}, Xiaotian Wang^{1,2}, Xinyue Huang⁴, Lei Li^{1,2}, Xiaolong Shan^{1,2}, Wanrou Hu^{1,2}, Yating Ming^{1,2}, Peng Wang^{1,2}, Jinming Xiao^{*1,3}, Huaifu Chen^{*1,2}, Xujun Duan^{*1,2}

1 The Clinical Hospital of Chengdu Brain Science Institute, School of Life Science and Technology, University of Electronic Science and Technology of China, Chengdu, 610054, PR China.

2 Brain-Computer Interface & Brain-Inspired Intelligence Key Laboratory of Sichuan Province, University of Electronic Science and Technology of China, Chengdu, 610054, PR China

3 Sichuan Provincial Center for Mental Health, Sichuan Provincial People's Hospital, University of Electronic Science and Technology of China, Chengdu, 610054, PR China.

4 Acupuncture and Tuina School, Chengdu University of Traditional Chinese Medicine, Chengdu, China.

*Corresponding author: Xujun Duan; Huaifu Chen ; Jinming Xiao

E-mail: duanxujun@uestc.edu.cn; chenhf@uestc.edu.cn; xiaojnming@gmail.com

Supplementary analysis

Development of a Network Diffusion Model

We model ASD progression as a diffusion process on a hypothesized brain network $\varsigma = \{\nu, \xi\}$ whose nodes $v_i \in \nu$ represent the i^{th} atypical ASD-related gray matter DEV, and the edges, $c_{i,j} \in \xi$, represent functional connectivity (FC) i and j . The structures v_i comes from DEV, and connection strength, $c_{i,j}$ is measured by FC¹. Consider an affected ($R2$) to an unaffected ($R1$) region. The number of diseased afferents from $R2$ to $R1$ is the product of the concentration, x_2 , of the disease factor and the inter-region connection strength, $c_{1,2}$. Conversely, there is a reverse diffusion from $R1$ to $R2$ proportional to $x_1 c_{2,1}$. The total concentration of the disease factor in $R1$ will therefore increase by $\beta(x_2 - x_1)c_{1,2}\delta t$ in a (short) instant δt , where β is the diffusivity constant controlling propagation speed. Assuming bidirectional pathways,

this leads, in the limit $\delta t \rightarrow 0$, to the following first-order differential equation:

$$\frac{dx_1}{dt} = |\beta c_{1,2}|(x_2 - x_1) \quad (1.1)$$

Spectral graph theory offers a sophisticated framework for extending Equation 1.1 to the entire network. Suppose the disease factor at time t at each node in the $x(t) = \{x(v, t), v \in \nu\}$. Under this formulation, Equation 1.1 can be generalized into the so-called "network heat equation," which describes the diffusion process across the entire system.

$$\frac{dx(t)}{dt} = -\beta Hx(t) \quad (1.2)$$

where H is the well-known graph Laplacian, with

$$H_{i,j} = \begin{cases} -c_{i,j} & \text{for } c_{i,j} \neq 0 \\ \sum_{i,j': e_{i,j'}} \in \mathcal{E} c_{i,j'} & \text{for } i = j \\ 0 & \text{otherwise} \end{cases} \quad (1.3)$$

This is the graph equivalent of the Laplacian diffusion operator, $\Delta x \triangleq \nabla^2 x$. Since our diffusion model represents brain network nodes as relatively large, anatomically distinct structures, the primary effect of localized transmission occurs within individual nodes. Given that the influence of local internode leakage is minimal, the network Laplacian H remains independent of self-connectivity within a node².

We hypothesize cortical atrophy in region k to be the accumulation of the disease process in k , modeled as the integral:

$$\phi_k(t) = \int_0^t x_k(\tau) d\tau \quad (1.4)$$

On the whole brain, this gives $\Phi(t) = \int_0^t x(\tau) d\tau$. From matrix algebra, Equation 2 is

satisfied by:

$$x(t) = e^{-\beta H t} x_0 \quad (1.5)$$

Where x_0 is the initial pattern of the disease process, on which the term $e^{-\beta H t}$ acts

essentially as a spatial and temporal blurring operator. We therefore call $e^{-\beta H t}$ the diffusion kernel, and Equation 1.5 is interpreted as the impulse response function of the network.

Divergence Measures: ED, TVD, and MMD

To evaluate the robustness of Kullback–Leibler (KL) divergence in capturing developmental distributional deviations, we incorporated three established distance metrics: Energy Distance (ED), Total Variation Distance (TVD), and Maximum Mean Discrepancy (MMD).

ED was computed following Székely and Rizzo ³, quantifying distributional divergence between ASD and TDC groups through pairwise Euclidean distances for each brain region. The DE between the independent random variables $X \sim P$ and $Y \sim Q$ is defined as:

$$ED(X, Y) = 2E|X - Y| - E|X - X'| - E|Y - Y'|$$

where $E|X| < \infty, E|Y| < \infty$, X' and Y' are independent copies of X and Y , respectively.

TVD is also known as the statistical difference, is a fundamental metric for measuring the difference between two probability distributions. For two distributions P and Q over the sample space Ω , their TVD is defined as ⁴:

$$\Delta TVD(P, Q) = \frac{1}{2} \sum_{x \in \Omega} |P(x) - Q(x)|$$

It corresponds to the L_1 -norm of the difference between P and Q , linking to other statistical divergences including relative entropy.

MMD was estimated using a Gaussian radial basis function (RBF) kernel. Given samples $X = \{x_1, x_2, \dots, x_m\}$ denotes the samples from distribution P and $Y = \{y_1, y_2, \dots, y_m\}$ denotes the samples from distribution Q . The unbiased estimate of squared MMD (MMD^2) is ⁵:

$$MMD^2(X, Y) = \frac{1}{m^2} \sum_{i,j=1}^m k(x_i, x_j) + \frac{1}{n^2} \sum_{i,j=1}^n k(y_i, y_j) - \frac{1}{mn} \sum_{i,j=1}^m k(x_i, y_j)$$

where $k(.,.)$ the Gaussian kernel with bandwidth chosen via the median heuristic.

Supplementary Figures

Supplementary Figure S1

The analysis of gray matter volume (GMV) DEV values across different age groups reveals distinct patterns of structural abnormalities in individuals with ASD. Throughout development, compared to TDCs, ASD individuals exhibit widespread regions with higher DEV values in early adolescence, while in late adolescence, they show widespread regions with lower DEV values. These findings suggest that ASD-related neurodevelopmental changes are dynamic and region-specific, highlighting potential disruptions in brain maturation trajectories that may underlie core ASD symptoms.

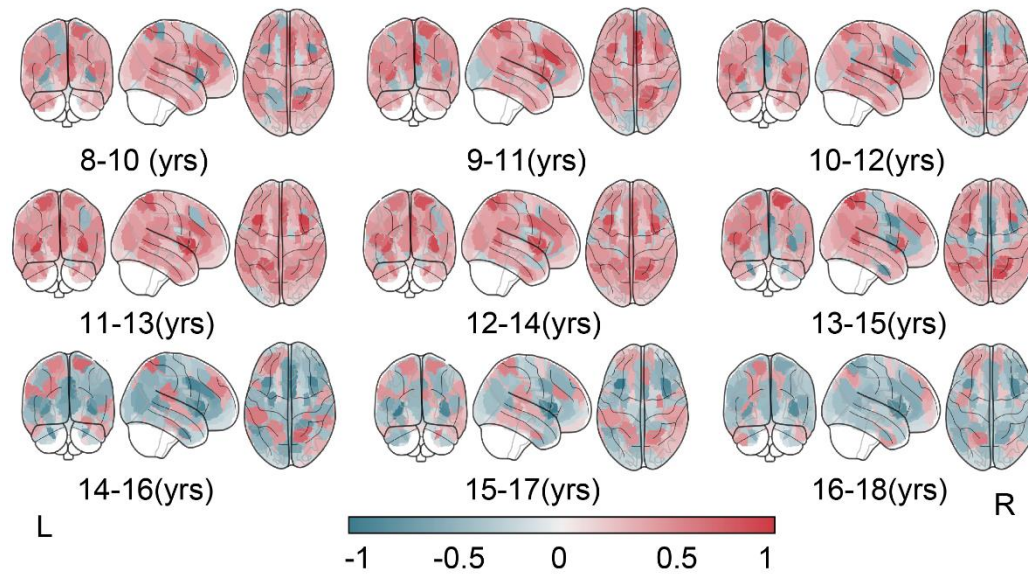


Fig. S1. The figure illustrates the DEV of GMV in individuals with ASD compared to TDCs across different age groups. Each subfigure represents the DEV within a specific age range, visualized using glass brain mapping. Red regions indicate higher DEV values in ASD individuals compared to TDCs, while blue regions indicate lower DEV values in ASD individuals. The color bar represents the value of DEV, with red colors indicating positive deviations and blue colors indicating negative deviations. yrs refers to the years.

Supplementary Figure S2-3

To evaluate the robustness of our findings to the choice of threshold, we conducted sensitivity analyses using alternative thresholds (top 5% and top 15%). Supplementary Figure S2 -3 demonstrate that the spatial distribution patterns of atypical regions remain largely consistent across thresholds. Although the spatial extent of identified regions varies slightly, the core regions exhibiting significant maturational asynchrony are preserved. Most importantly, the key findings regarding maturational transition patterns are robust across these different thresholds.

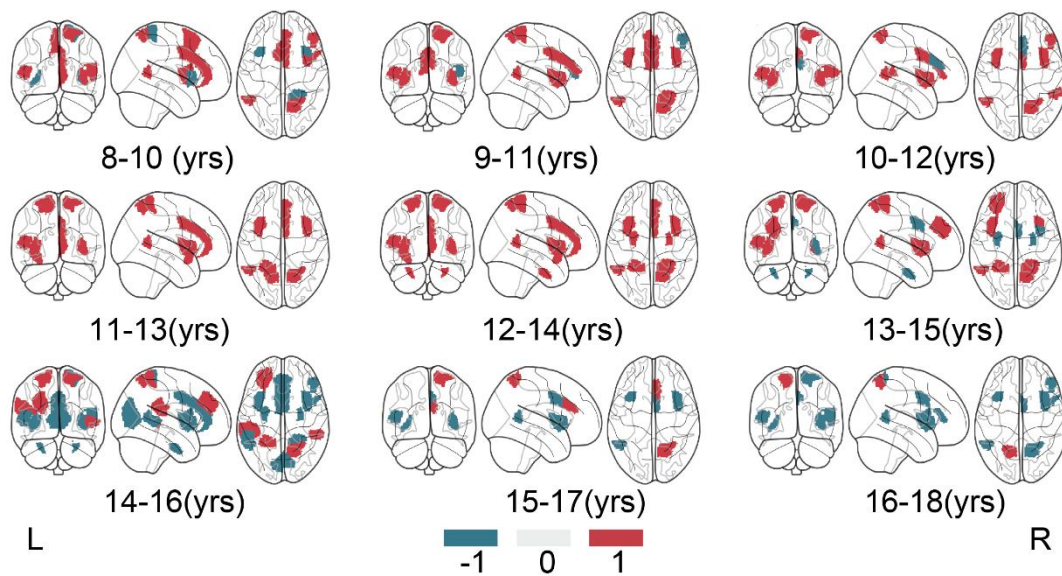


Fig. S2. Glass brain maps illustrating the values in the top 5% of spatial distribution deviation (DEV). Red areas indicate regions where DEV values are higher in the ASD group compared to TDCs, with the color bar label 1 representing the top 0–5% of positive deviations. Blue regions represent areas where DEV values are lower in the ASD group, with the label –1 corresponding to the top 0–5% of negative deviations. Label 0 denotes regions with DEV values outside these thresholds, darker colors indicate larger DEV scores. yrs refers to the year.

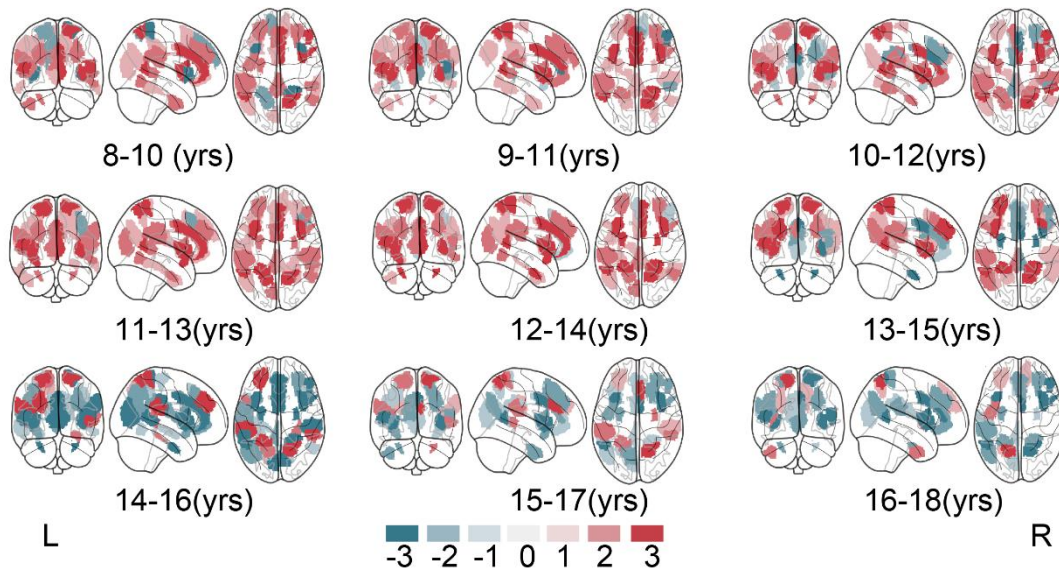


Fig. S3. Glass brain maps illustrating the values in the top 15% of spatial distribution deviation (DEV). Red areas indicate regions where DEV values are higher in the ASD group compared to TDCs (with color bar label 1 representing the top 11-15% of positive DEV values, and label 2 representing the top 6-10%, and label 3 representing the top 0-5%). Blue areas indicate regions where DEV values are lower in the ASD group (with color bar label 1 representing the top 11-15% of positive DEV values, and label 2 representing the top 6-10%, and label 3 representing the top 0-5%). Label 0 denotes regions with DEV values outside these thresholds, darker colors indicate larger DEV scores. yrs refers to the years.

Supplementary Figure S4

Fig. S4. shows the nine windows, where a significant correlation was observed between the nodal DEV and the mean of its directly connected neighbors.

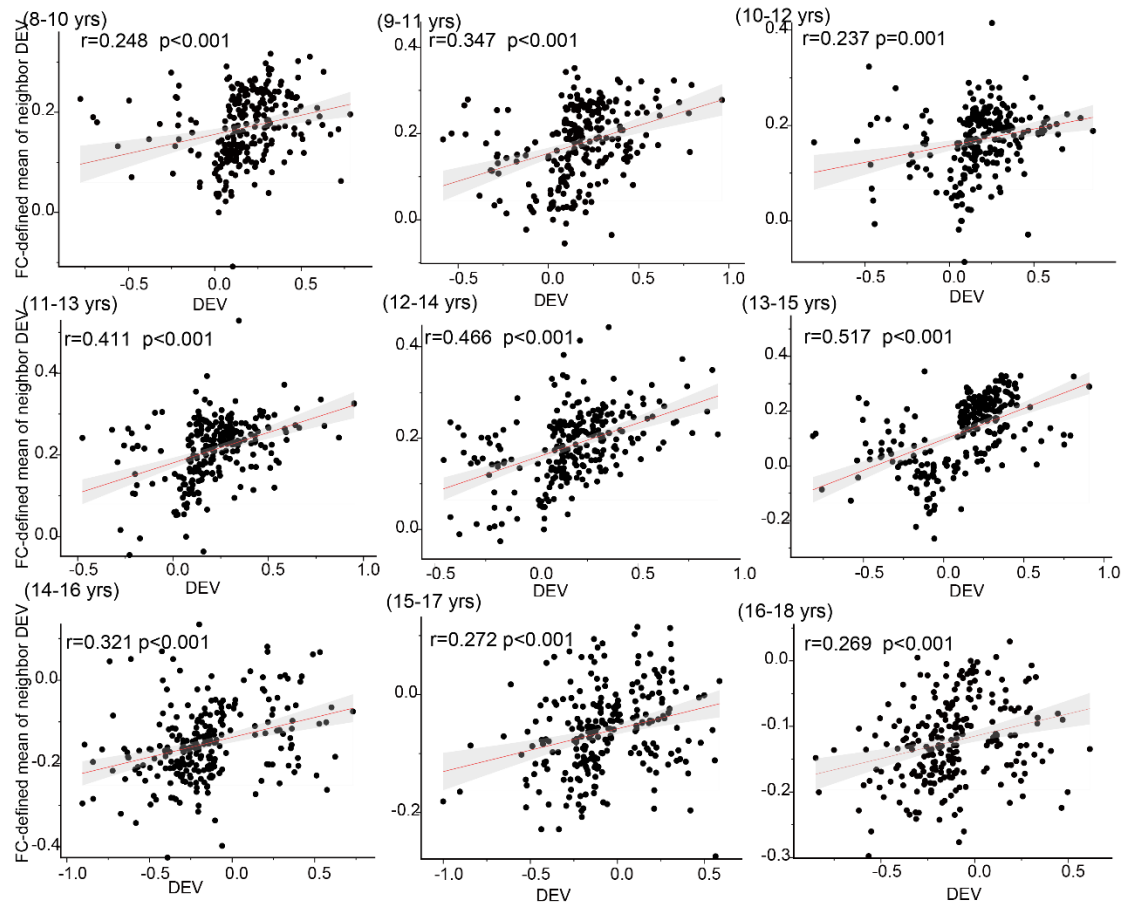


Fig. S4. A significant correlation was observed between the nodal distribution deviation (DEV) and the mean of its directly connected neighbors across the nine windows (Pearson correlation, two-sided test). The scatter plot illustrates the relationship between individual node-level DEV values (x-axis) and the average DEV values of their directly connected neighbors (y-axis). Each dot represents one brain region. The solid line denotes the linear fit, and the shaded area indicates the 95% confidence interval. yrs refers to the years.

Supplementary Figure S5

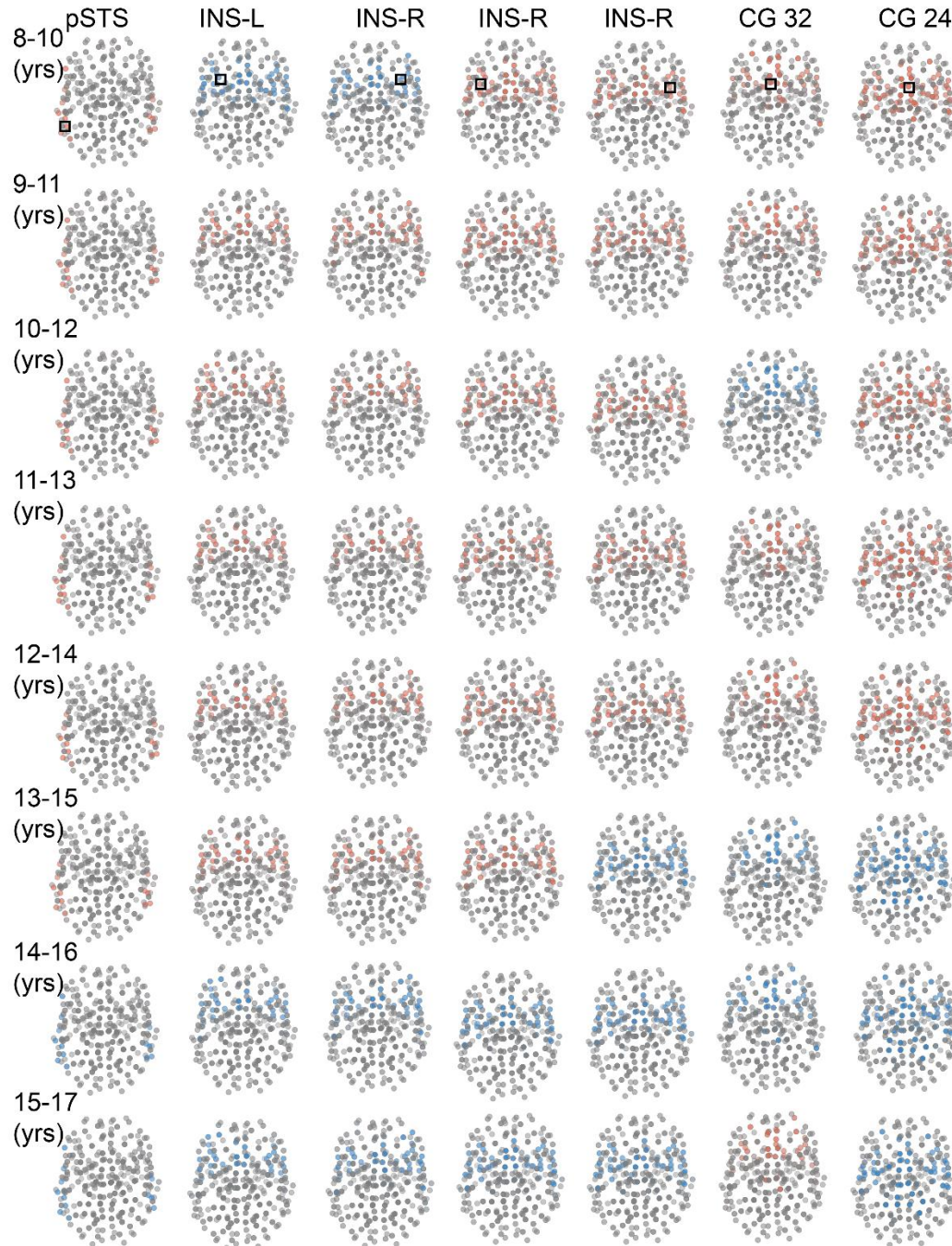


Fig. S5. Spatiotemporal diffusion patterns of seven representative nodes across all age group. Red and blue nodes denote regions with positive and negative DEV values, respectively. Nodes boxed in black indicate key regions referenced in seven representative nodes. yrs refers to the years.

Supplementary Figure S6

In this study, to further validate the reliability of the Kullback-Leibler (KL) divergence computation model, we introduced three commonly used probability distribution distance metrics for comparative analysis, including Energy Distance (ED) (Fig. S6A), Maximum Mean Discrepancy (MMD) (Fig. S6B), and Total Variation Distance (TVD) (Fig. S6C). Energy Distance is a metric based on Euclidean distance that effectively quantifies the overall discrepancy between two distributions. Total Variation Distance, on the other hand, captures the maximum deviation between probability mass functions, reflecting local variations in distributions. Maximum Mean Discrepancy is a non-parametric kernel-based metric that maps probability distributions into a reproducing kernel Hilbert space to measure their differences.

In the current study, we computed the correlation between KL divergence and each of these three metrics and conducted statistical significance testing using permutation tests. The results indicated that ED, TVD, and MMD were all highly correlated with KL divergence ($p < 0.001$, FDR-corrected), demonstrating a high degree of consistency among these distance measures in quantifying probability distribution differences. These findings further validate the rationality of the KL divergence model and support its effectiveness as a robust metric for measuring distributional discrepancies.

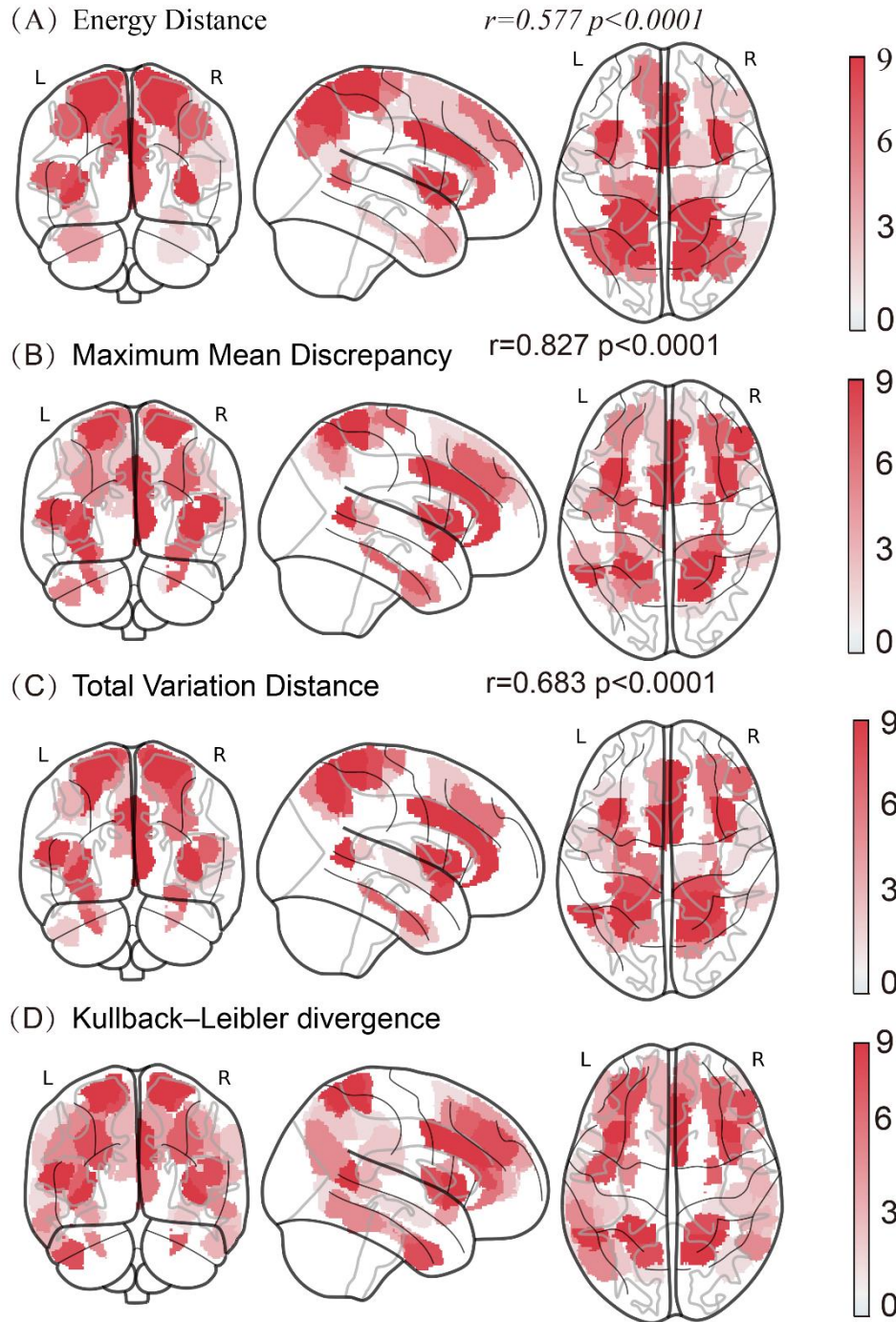


Fig. S6. Validation of the KL divergence computation model using three commonly used probability distribution distance metrics. (A) The glass brain map shows the atypical regions identified using Energy Distance, with a correlation of $r = 0.576$ ($p < 0.001$) with the brain regions identified by KL divergence. **(B)** The glass brain map shows the atypical regions identified using Maximum Mean Discrepancy; with a correlation of $r = 0.827$ ($p < 0.001$) with the brain regions identified by KL divergence. **(C)** The glass brain map shows the atypical regions identified using Total Variation Distance, with a correlation of $r = 0.683$ ($p < 0.001$) with the brain regions identified

by KL divergence; **(D)** The glass brain map shows the atypical regions identified using KL divergence. The color bar represents the frequency with which each region was identified as atypical across age groups.

Supplementary Figure S7

Fig. S7. illustrates the atypical brain regions identified using both 200-node and 400-node parcellation schemes. To assess the impact of parcellation resolution on our findings, we conducted analyses with these two distinct brain atlases and compared the spatial distributions of the identified atypical regions. The results showed no significant differences between the three resolutions, indicating that the observed atypical patterns are consistent and not dependent on the specific choice of parcellation. This consistency underscores the methodological reliability of our approach and reinforces the robustness of our findings across different spatial scales.

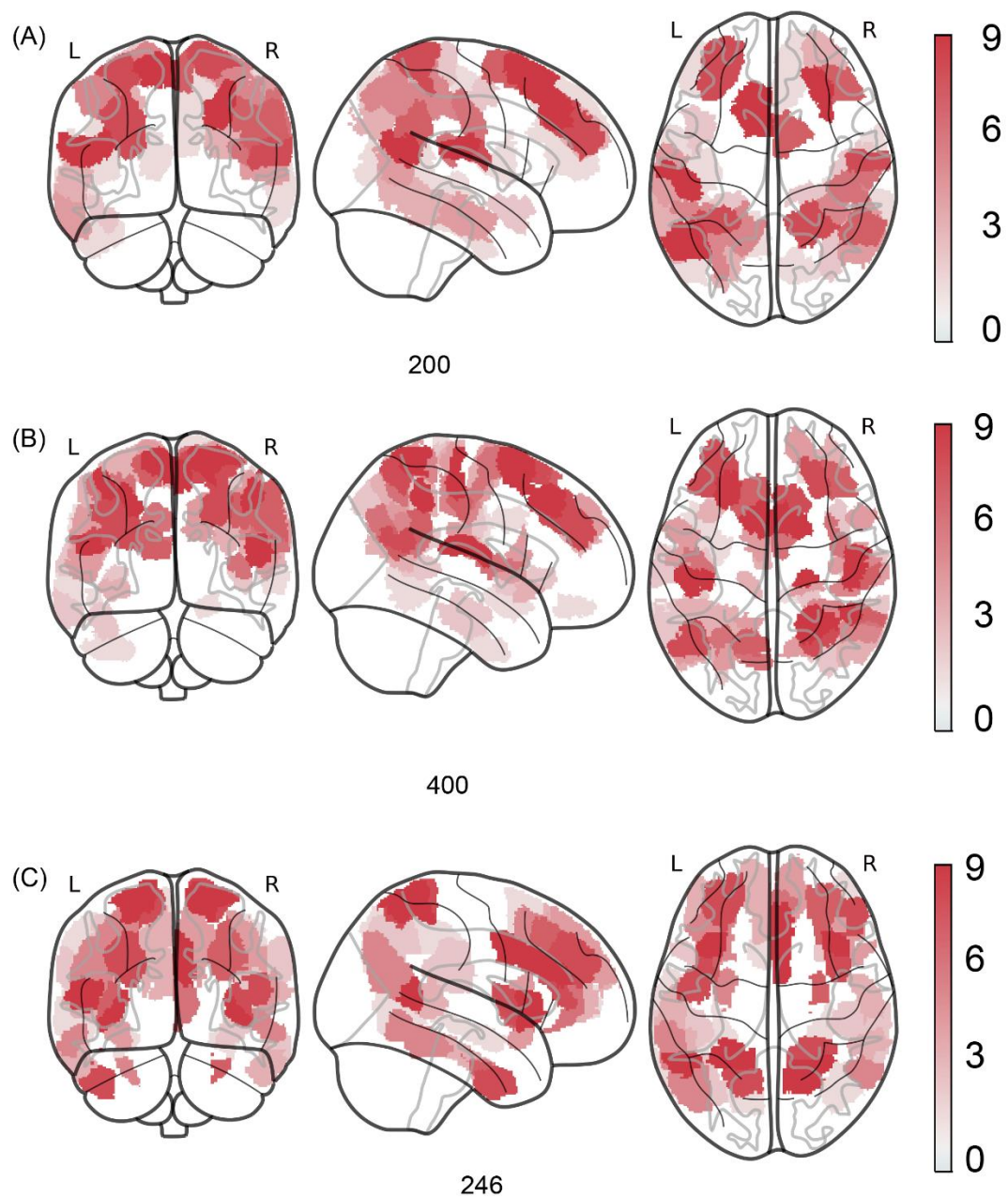


Fig. S7. Atypical brain regions identified using 200-node and 400-node parcellation resolutions. (A) The glass brain map shows the atypical brain regions identified using the 200-node brain atlas. (B) The glass brain map shows the atypical brain regions identified using the 400-node brain atlas. (C) The glass brain map shows the atypical brain regions identified using the 246-node brain atlas. The color bar reflects the frequency with which each brain region was classified as atypical across sliding age windows. The color bar represents the frequency with which each region was identified as atypical across age groups.

Supplementary Figure S8

To assess the stability and robustness of our findings, we conducted additional analyses by examining specific subgroups within our dataset. First, we computed the results using only male participants from a single brain region and found a strong correlation with the main findings reported in the manuscript ($r = 0.963$, $p < 0.0001$), indicating a high degree of consistency. Additionally, we performed a separate analysis using data from the largest data center in our study, the New York University (NYU) dataset, and observed a correlation of ($r = 0.725$, $p < 0.0001$) with the primary results. These findings suggest that our results are reproducible across different subgroups and data sources, further supporting the reliability of our methodology and the generalizability of the observed patterns.

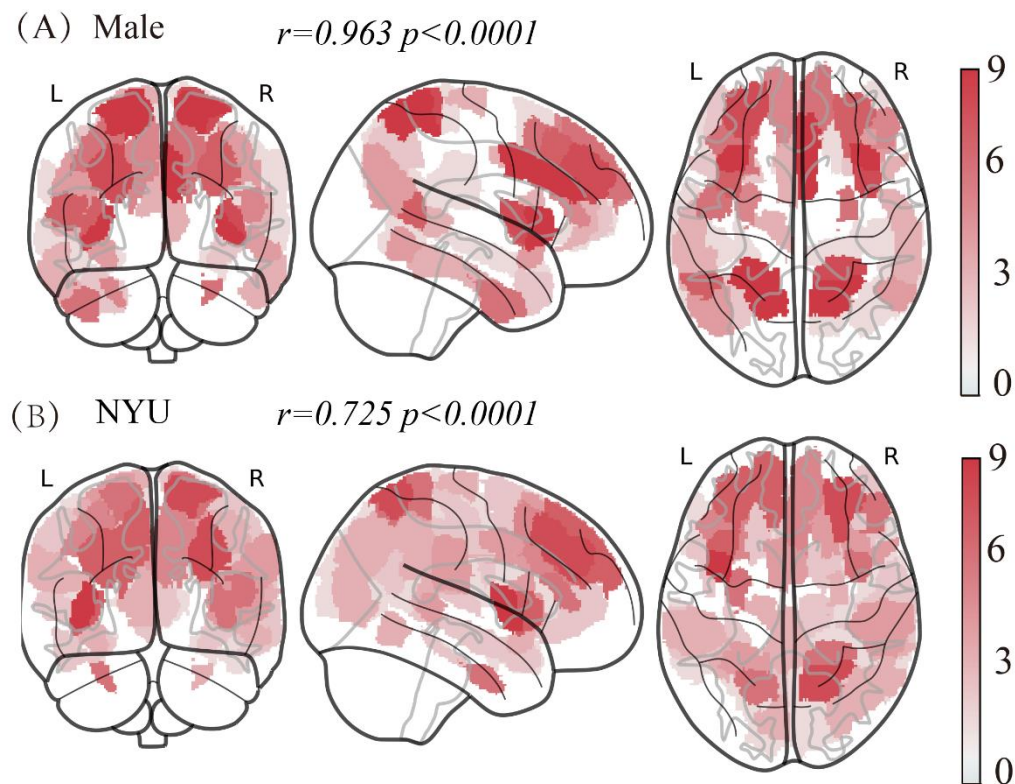


Fig. S8. Stability and robustness of the findings across different subgroups and data sources.

(A) The glass brain map shows the atypical brain regions of male participants, with a correlation of $r = 0.963$ ($p < 0.001$) with the atypical regions reported in the manuscript. **(B)** The glass brain map shows the atypical brain regions of NYU participants, with a correlation of $r = 0.725$ ($p < 0.001$)

with the atypical regions reported in the manuscript. The color bar represents the frequency with which each region was identified as atypical across age groups.

Supplementary Figure S9

Fig. S9. shows three randomly generated probability density functions (PDFs) with the same standard deviation: PDF 1, PDF 2, and PDF 3. PDF 2 is represented by a solid line, while PDF 1 and PDF 3 are shown as dashed lines, with PDF 1 positioned to the left of PDF 2 and PDF 3 to the right. As illustrated, the Kullback–Leibler (KL) divergence between the dashed PDFs (PDF 1 and PDF 3) and PDF 2 yields only positive values regardless of the direction of shift, due to its inherent asymmetry and non-negativity. Therefore, KL divergence alone cannot determine whether a distribution is shifted to the left or right relative to the reference distribution.

To address this limitation, we incorporate the concept of the mathematical expectation, defined for a continuous probability distribution $f(x)$ as:

$$E[x] = \int_{-\infty}^{\infty} xf(x)dx$$

We compute the expected values of both the dashed PDFs and PDF 2. By taking the difference between them ($E[PDF_{dashed}] - E[PDF 2]$), we derive a directional indicator. A negative value indicates that the dashed PDF is shifted to the left of the reference PDF 2, while a positive value indicates a shift to the right.

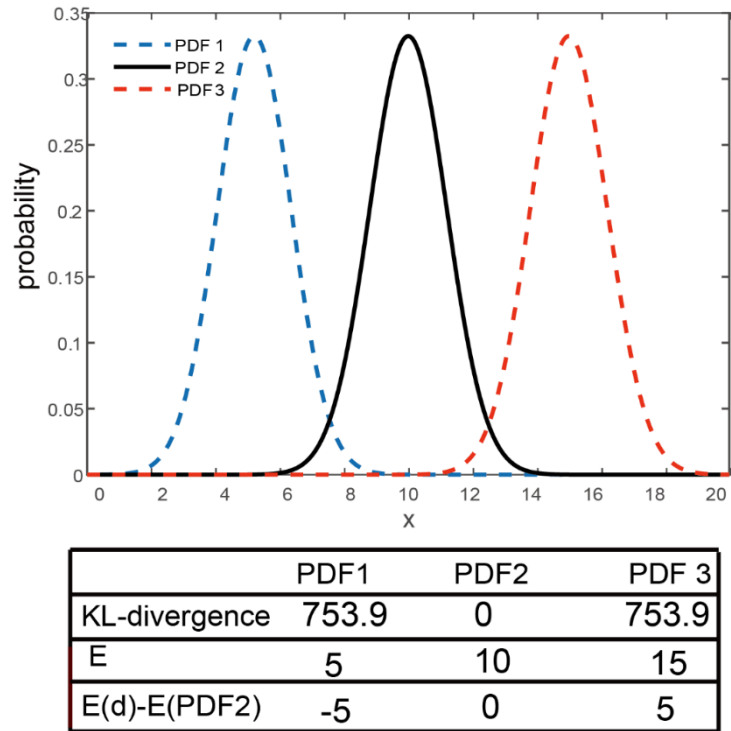


Fig. S9. PDFs and their corresponding KL-divergence values. The plot illustrates three PDFs: PDF 1 (blue dashed line), PDF 2 (black solid line), and PDF 3 (red dashed line). The accompanying table presents the KL- divergence values of each PDF relative to PDF 2, along with their respective expected values (E). The last row shows the difference between the expected value of each PDF and that of PDF 2.

Supplementary Figure S10

NDM model performance

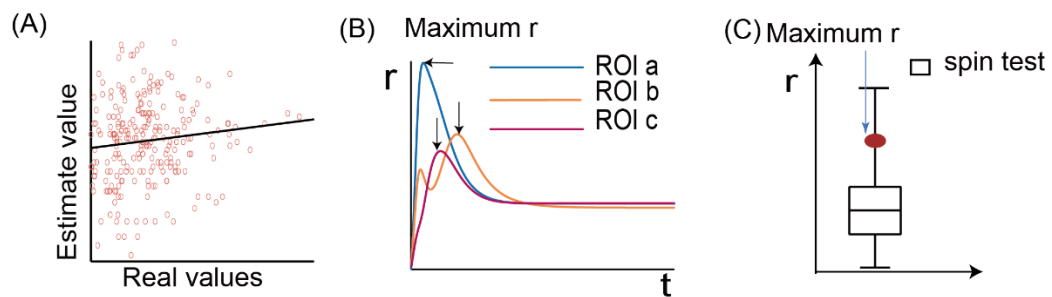


Fig. S10. Evaluation of network diffusion model (NDM) performance. (A): The scatter plot shows the relationship between observed and estimated distribution deviation (DEV) values across brain regions, with the strength of association quantified using Pearson correlation (r). (B): For each seed region (e.g., ROI a, b, c), the correlation (r) between observed and simulated DEV values was

computed across diffusion time points. The maximum *r* value was extracted to characterize the optimal model fit. **(C)**: The observed maximum *r* value (red dot) was compared against a null distribution of maximum *r* values generated from spatially permuted (spin test) surrogates. Statistical significance was determined by assessing whether the observed value exceeded the 95th percentile of the null distribution.

Supplementary Table

Supplementary Table S1-3

Table S1 summarizes the mean and standard deviation of DEV values in atypical brain regions for each age group, stratified by sliding age windows. These values provide a quantitative summary of the developmental shift in deviation patterns. Specifically, the mean DEV values were positive to the 13–15 age group and subsequently became negative, indicating a trajectory in which deviation values shift from early positive to later negative with increasing age. In addition, we also calculated the mean $\pm$ SD of DEV values within atypical regions defined by alternative thresholds (top 5% and top 15%) (Table S2 and S3). These analyses revealed a consistent pattern: mean DEV values were positive in earlier age group (before 13–15 age group) and turned negative in later stages, reinforcing the observed developmental transition. These analyses revealed a consistent pattern: mean DEV values were positive in earlier age windows and turned negative in later stages, reinforcing the observed developmental transition.

Table S1. Mean and standard deviation of atypical regional (top 10%) DEV values

<i>Age (yrs)</i>	<i>8-10</i>	<i>9-11</i>	<i>10-12</i>	<i>11-13</i>	<i>12-14</i>	<i>13-15</i>	<i>14-16</i>	<i>15-17</i>	<i>16-18</i>
<i>Mean</i>	<i>0.032</i>	<i>0.044</i>	<i>0.037</i>	<i>0.060</i>	<i>0.056</i>	<i>0.001</i>	<i>-0.051</i>	<i>-0.020</i>	<i>-0.034</i>
<i>SD</i>	<i>0.189</i>	<i>0.180</i>	<i>0.185</i>	<i>0.193</i>	<i>0.184</i>	<i>0.197</i>	<i>0.240</i>	<i>0.175</i>	<i>0.165</i>

Note: SD refers to the standard; yrs refers to the years.

Table S2. Mean and standard deviation of atypical regional (top 5%) DEV values

<i>Age (yrs)</i>	<i>8-10</i>	<i>9-11</i>	<i>10-12</i>	<i>11-13</i>	<i>12-14</i>	<i>13-15</i>	<i>14-16</i>	<i>15-17</i>	<i>16-18</i>
<i>Mean</i>	<i>0.014</i>	<i>0.029</i>	<i>0.021</i>	<i>0.035</i>	<i>0.040</i>	<i>0.012</i>	<i>-0.039</i>	<i>-0.012</i>	<i>-0.020</i>
<i>SD</i>	<i>0.145</i>	<i>0.153</i>	<i>0.137</i>	<i>0.159</i>	<i>0.164</i>	<i>0.164</i>	<i>0.209</i>	<i>0.127</i>	<i>0.128</i>

Note: SD refers to the standard; yrs refers to the years.

Table S3. Mean and standard deviation of atypical regional (top 15%) DEV values

<i>Age (yrs)</i>	<i>8-10</i>	<i>9-11</i>	<i>10-12</i>	<i>11-13</i>	<i>12-14</i>	<i>13-15</i>	<i>14-16</i>	<i>15-17</i>	<i>16-18</i>
<i>Mean</i>	<i>0.048</i>	<i>0.061</i>	<i>0.043</i>	<i>0.088</i>	<i>0.068</i>	<i>0.024</i>	<i>-0.064</i>	<i>-0.035</i>	<i>-0.036</i>
<i>SD</i>	<i>0.203</i>	<i>0.207</i>	<i>0.207</i>	<i>0.213</i>	<i>0.206</i>	<i>0.218</i>	<i>0.254</i>	<i>0.198</i>	<i>0.184</i>

Note: SD refers to the standard; yrs refers to the years.

Supplementary Table S4

Table S4 presents the site information and age distribution of all participants.

Table S4. Information of site

Site\age(yrs)	8	9	10	11	12	13	14	15	16	17	Total
YALE (ASD/TDC)	1/1	1/2	0/2	3/0	2/1	3/2	2/2	0/2	4/1	0/1	16/14
USM	--	0/1	--	1/0	1/1	0/2	0/1	2/0	3/1	2/1	9/7
UM	0/1	4/2	1/4	2/5	4/1	7/3	5/2	4/8	6/3	1/12	34/41
UCLA	1/1	3/4	5/1	4/6	3/6	4/6	5/5	2/3	3/0	3/1	33/33
TRINITY	--	--	--	--	0/2	3/2	1/4	3/3	0/1	1/1	8/13
STANFORD	4/2	2/3	1/1	3/1	1/1	--	--	--	--	--	11/8
SDSU	0/1	3/2	5/1	1/2	2/3	4/3	4/5	4/3	5/6	7/3	35/29
PITT	--	--	--	1/0	1/3	1/0	2/1	0/2	2/0	0/1	7/7
OLIN	--	--	--	--	2/0	--	1/1	1/1	1/1	1/1	6/4
NYU	12/11	7/6	10/9	8/11	5/6	3/3	5/5	3/4	1/6	1/1	55/62
LEUVEN	--	--	--	--	2/4	2/0	6/3	1/3	0/4	1/0	12/14
KKI	6/11	1/24	5/20	3/13	1/8	--	--	--	--	--	16/76

CALTECH	--	--	--	--	0/1	--	--	--	--	1/1	1/2
UCD	--	--	--	--	1/1	3/1	2/1	1/2	3/2	3/2	13/9
TCD	--	--	2/0	--	0/2	0/2	1/1	0/3	0/1	0/1	3/10
SU	1/0	3/4	4/2	3/3	5/4	0/1	--	--	--	--	16/14
IU	--	--	--	--	--	--	--	--	--	2/0	2/0
GU	1/9	3/4	6/8	6/5	5/4	3/2	--	--	--	--	24/32
Total	26/37	27/52	39/48	35/46	35/48	33/27	34/31	21/34	28/26	23/26	301/375

Note: yrs refers to the years.

Supplementary Table S5

This study systematically evaluated the statistical stability of functional connectivity (FC) networks across different sparsity thresholds. The experimental results (see Table S5) indicate that within the sparsity range of 10% to 50%, the correlation coefficients (r -values) between adjacent windows were predominantly significant (most p -values were less than 0.001), demonstrating the robustness of FC network features within this range. In the manuscript, we present the results corresponding to the 10% sparsity threshold.

Table S5. Correlation coefficients and corresponding p -values for different neighboring windows across various percentages.

sparsity levels		Age group (yrs)							
		8-10-9-11	9-10-10-12	10-12-11-13	11-13-12-14	12-14-13-15	13-15-14-16	14-16-15-17	15-17-16-18
10%	r	0.560	0.538	0.623	0.587	0.642	0.722	0.642	0.669
	p	all<0.001							
20%	r	0.697	0.623	0.663	0.627	0.699	0.580	0.694	0.587
	p	all<0.001							
30%	r	0.668	0.659	0.665	0.619	0.670	0.564	0.703	0.565
	p	all<0.001							
40%	r	0.670	0.663	0.496	0.602	0.651	0.577	0.709	0.573

	<i>p</i>	<0.001	<0.001	0.041	<0.001	<0.001	<0.001	<0.001	<0.001
50%	<i>r</i>	0.663	0.676	0.470	0.526	0.5820	0.064	0.684	0.579
	<i>p</i>	<0.001	<0.001	<0.064	=0.001	<0.001	<0.001	<0.001	<0.001

Notes: p represents the p-value obtained from the permutation test; yrs refers to the years.

Supplementary Table S6

Table S6 presents the demographic characteristics of participants from the NYU site. The sample included 55 individuals with ASD (11.35 ± 2.45 years) and 62 TDCs (11.92 ± 2.68 years).

Table S6. NYU subject demographics

	ASD (55)	TDCs (62)	<i>p</i>
Age (yrs)	11.35±2.45	11.92±2.68	0.23 ^a
Sex(male/fmale)	49/6	54/8	0.74 ^b
Handedness(Right/left/mixed/)	43/8/4	58/4/0	0.02 ^b *
FIQ	106.45±18.18	114.24±15.44	0.01 ^a *

Mean and standard deviation are given for each continues measure; yrs refers to the years.

a The P value was obtained by a 2-sample t-test, 2-tailed.

b The P value was obtained by the Pearson's chi-square test.

* Significant at $p < 0.05$.

Supplementary Materials

Supplementary Data 1

Supplementary Data 2

Supplementary Data 3

Supplementary Data 4

Reference

- Behrens, T. E., Berg, H. J., Jbabdi, S., Rushworth, M. F. & Woolrich, M. W. Probabilistic diffusion tractography with multiple fibre orientations: What can we gain? *neuroimage* **34**, 144-155 (2007).
- Raj, A., Kuceyeski, A. & Weiner, M. A network diffusion model of disease progression in dementia. *Neuron* **73**, 1204-1215, doi:10.1016/j.neuron.2011.12.040 (2012).
- Székely, G. J. & Rizzo, M. L. Energy statistics: A class of statistics based on distances. *Journal*

of Statistical Planning and Inference **143**, 1249-1272, doi:10.1016/j.jspi.2013.03.018 (2013).

- 4 Feng, W., Liu, L. & Liu, T. in *Proceedings of the 2024 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*. 1766-1791 (SIAM).
- 5 Jia, X., Zhao, M., Di, Y., Yang, Q. & Lee, J. Assessment of Data Suitability for Machine Prognosis Using Maximum Mean Discrepancy. *IEEE Transactions on Industrial Electronics* **65**, 5872-5881, doi:10.1109/tie.2017.2777383 (2018).