The Neuro-Developing Engine: A Surrogate Longitudinal Model for Infant Brain Dynamics via Causal Distillation

Dan Hu, Jiale Cheng, Kangfu Han, Zhengwang Wu, Li Wang, Weili Lin, Gang Li✉

Department of Radiology and Biomedical Research Imaging Center,
University of North Carolina at Chapel Hill, Chapel Hill, NC, 27599, USA
gang_li@med.unc.edu

Abstract. Understanding how individual infant brain function evolves over development requires models that capture personalized longitudinal changes in functional dynamics, rather than static connectivity estimates. However, existing data-driven approaches largely focus on longitudinal functional connectivity (FC) prediction, leaving a significant gap in mechanistic interpretability, causal analysis, and grounded support for targeted intervention in atypical brain development. To address this challenge, we propose the first Neuro-Developing Engine (NDE), a novel surrogate longitudinal model of infant brain dynamics utilizing resting-state functional MRI (rs-fMRI) data. By learning age-dependent transitions between individualized surrogate brains at the individual level, NDE not only predicts the evolution of surrogate brain dynamics across continuous ages but also identifies brain regions exhibiting strong modulatory effects during development. Specifically, 1) leveraging individualized Neural Perturbational Inference (NPI) surrogate models that capture subject-specific brain dynamics, we design a developmental residual model equipped with FiLM-conditioned generators to distill explicit age-driven modulation between younger and older brain dynamics, conditioned on graph-embedded individual identity and ages. 2) To enhance causal and age-sensitive representations, we combine causal distillation and Jacobian spectrum objectives with a soft-label contrastive loss that enforces alignment between predicted residual dynamics and age gaps. 3) To enable inverse-control analysis, we introduce region-level "counterfactual developmental knobs" that modulate residual dynamics, allowing the quantification of how age-dependent perturbations reshape FC over development. Extensive experiments demonstrate NDE's ability to model individualized developmental dynamics, predict connectivity changes, and support counterfactual inverse-control simulation, highlighting its potential for mechanistic understanding and guiding early intervention in typical and atypical neurodevelopment.

Keywords: Infant Brain Development, Surrogate Brain Models, Resting-state fMRI (rs-fMRI).

1 Introduction

Early neurodevelopment represents the most dynamic phase of human brain maturation, characterized by rapid functional reorganization that shapes lifelong cognitive,

emotional, and behavioral trajectories [1-3]. Capturing these highly individualized developmental dynamics with resting-state functional MRI (rs-fMRI) is essential for identifying early biomarkers of both typical and atypical neurodevelopment [4-6]. Despite substantial progress in mapping early functional patterns, most data-driven approaches remain confined to static connectivity estimates, such as descriptive group-level trajectories [2] or longitudinal functional connectivity (FC) prediction [7, 8]. While informative for developmental trends, these methods are vulnerable to scan noise [9, 10] and offer limited mechanistic insight, leaving a critical gap in explaining how specific alterations in neural dynamics reshape developmental outcomes and limiting our ability to identify trajectory divergence or guide targeted early intervention.

Neural Perturbational Inference (NPI) models [11] derived from rs-fMRI offer a promising step toward mechanistic modeling by enabling in silico causal exploration of interactions among brain regions. By training artificial neural networks (ANNs) as subject-specific neural surrogates that capture functional dynamics, NPI advances digital brain modeling from mechanism-driven brain structural frameworks [12, 13] to data-driven functional surrogate modeling. However, current NPI formulations are restricted to a single scan time, capturing only static dynamics without explicitly modeling developmental transitions, and therefore cannot represent continuous developmental evolution or mechanistically interrogate changes in brain dynamics over time.

To bridge this gap, we propose the Neuro-Developing Engine (NDE), the first rs-fMRI-based surrogate longitudinal framework for modeling infant brain dynamics. Rather than predicting static connectivity, NDE learns developmental transitions between individualized surrogate brains, enabling mechanistic simulation of how functional dynamics evolve over ages. Our main contributions are summarized as follows:

– We introduce NDE, a novel framework that extends Neural Perturbational Inference from single-timepoint surrogate modeling to continuous developmental trajectory modeling, enabling longitudinal simulation of individualized brain dynamics.

– We propose a Feature-wise Linear Modulation (FiLM)-conditioned residual transition module that distills age-driven evolution between surrogate brains by embedding identity and age into dynamic transitions, while integrating causal distillation, Jacobian spectrum regularization, and soft-label contrastive learning to align residual dynamics with developmental progression.

– We introduce brain-region-level counterfactual developmental knobs for in silico analysis of how localized neural modulation reshapes future FC trajectories.

Experiments on typically and atypically developing infants demonstrate that NDE effectively models individualized developmental dynamics, detects abnormal connectivity evolution, and supports interpretable counterfactual simulations.

2 Method

2.1 Model Description

The framework of our proposed Neuro-Developing Engine (NDE) is depicted in **Fig. 1** and detailed below. It consists of individual surrogate brain learning, developmental residual transition module, trajectory divergence identification, and counterfactual

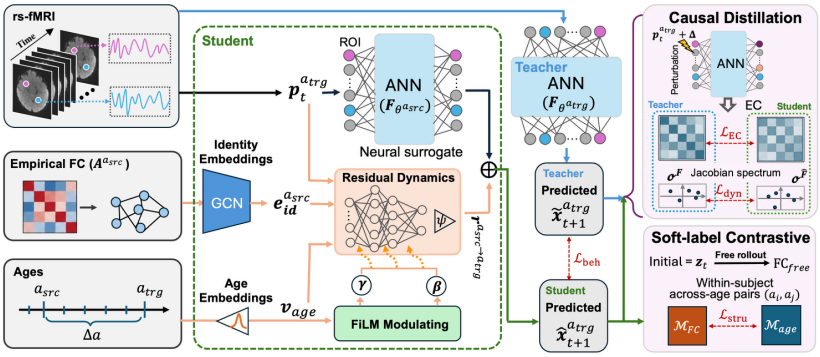

Fig. 1. The framework of our proposed NDE model.

developmental knobs (i.e., brain regions).

Given an rs-fMRI scan from subject s, $\boldsymbol{X}^{s,a} = \{\boldsymbol{x}_t^{s,a}\}_{t=1}^{T}$, where $\boldsymbol{x}_t^{s,a} \in \mathbb{R}^K$ denotes the BOLD signals of K brain regions at temporal frame t, acquired at chronological age a and T is the total number of frames in the scan.

Individual surrogate brain learning. For each subject-age pair (s, a), a subject-specific neural surrogate [11] is learned to model functional dynamics from the observed BOLD signals, which is formulated as a state-evolution process learned from $\boldsymbol{X}^{s,a}$. An artificial neural network (ANN) $\boldsymbol{F}_{\theta^{s,a}}$ approximates the temporal changes by predicting the next brain state $\boldsymbol{x}_{t+1}^{s,a}$ from a context window of its previous ℓ frames, $\boldsymbol{p}_t^{s,a} = \left[\boldsymbol{x}_{t-\ell+1}^{s,a}, \cdots, \boldsymbol{x}_t^{s,a}\right]$, $\boldsymbol{x}_{t+1}^{s,a} = \boldsymbol{F}_{\theta^{s,a}}(\boldsymbol{p}_t^{s,a}) + \boldsymbol{\epsilon}_t$, where $\theta^{s,a}$ denotes (s, a)-specific ANN parameters and $\boldsymbol{\epsilon}_t$ models observation noise. The surrogate is optimized through

$$\mathcal{L}_{ANN} = \sum_{t=\ell}^{T-1} \|\boldsymbol{x}_{t+1}^{s,a} - \boldsymbol{F}_{\theta^{s,a}}(\boldsymbol{p}_t^{s,a})\|_2^2 \quad (1)$$

Each scan is thus represented as a differentiable surrogate capturing individualized functional dynamics, forming the basis for developmental transition modeling. The learned surrogate enables estimation of effective connectivity (EC), characterizing directed causal influences between brain regions. EC is computed by quantifying how perturbations in one region influence the predicted future state of another, approximated via the Jacobian, $\boldsymbol{J}_t^{s,a} = \partial \boldsymbol{F}_{\theta^{s,a}}(\boldsymbol{p}_t^{s,a}) / \partial \boldsymbol{p}_t^{s,a}$, whose temporal average yields the EC estimate. The subject index s is omitted hereafter when no ambiguity arises.

Developmental residual transition module. To model the longitudinal evolution of brain dynamics, we introduce a developmental residual transition module that learns continuous age-dependent transitions between surrogates. Rather than training independent age-specific surrogates, this module predicts the developmental shift from source age a_{src} to target age a_{trg} while preserving subject-specific individuality. It comprises three components: a graph-based identity encoder, age embedding, and an age-FiLM residual dynamics module fortified by an age-gap-aware gating mechanism.

Subject identity encoder. To separate age-related changes from intrinsic subject-specific functional architecture, we extract an identity embedding $\boldsymbol{e}_{id}$ using a Graph

Convolutional Network (GCN) applied to the source-age empirical functional connectivity (FC) matrix $\boldsymbol{A}^{a_{src}} \in \mathbb{R}^{K\times K}$. Node features are initialized as $\boldsymbol{H}^{(0)} = \boldsymbol{A}^{a_{src}}$ and propagated through L GCN layers, followed by global mean pooling:

$$\boldsymbol{e}_{id} = 1/K \sum_{k=1}^{K} \boldsymbol{H}_k^{(L)} \tag{2}$$

The resulting embedding captures subject-specific connectome topology and establishes a stable identity prior for developmental modeling.

High-dimensional age embedding. Original scalar ages offer limited representational capacity for modeling nonlinear and heterogeneous developmental dynamics. To address this, chronological ages a_{src}, a_{trg}, and their age gap $\Delta a = a_{trg} - a_{src}$ are mapped into a high-dimensional feature space using Fourier feature encoding $\phi(\cdot)$ to capture smooth trends. The resulting age embeddings are:

$$\boldsymbol{v}_{age} = [\phi(a_{src})||\phi(a_{trg})||\phi(\Delta a)] \tag{3}$$

This high-frequency temporal encoding enriches age representation and enables the model to parameterize continuous, nonlinear developmental trajectories.

Age-FiLM residual dynamics module. Neurodevelopment largely reflects gradual refinement of existing functional circuitry rather than abrupt reorganization. Instead of learning an independent new model $\boldsymbol{F}_{\theta^{a_{trg}}}$, we adapt the source-age surrogate $\boldsymbol{F}_{\theta^{a_{src}}}$ with a developmental residual that realizes dynamic transition across ages while preserving subject-specific structure. Given neural state $\boldsymbol{p}_t^{a_{trg}}$, identity embedding $\boldsymbol{e}_{\boldsymbol{id}}^{a_{src}}$, and age embedding $\boldsymbol{v}_{age}$, the model predicts a residual $\boldsymbol{r}^{a_{src}\to a_{trg}} \in \mathbb{R}^K$:

$$\hat{\boldsymbol{x}}_{t+1}^{a_{trg}} = \hat{\boldsymbol{F}}_{\theta^{a_{trg}}}(\boldsymbol{p}_t^{a_{trg}}) = \boldsymbol{F}_{\theta^{a_{src}}}(\boldsymbol{p}_t^{a_{trg}}) + \boldsymbol{r}^{a_{src}\to a_{trg}} \tag{4}$$

The residual $\boldsymbol{r}^{a_{src}\to a_{trg}}$ is generated by an age-FiLM [14] modulated network. A fused representation $\boldsymbol{h}^{(0)} = \boldsymbol{W}_{in}[\boldsymbol{p}_t^{a_{trg}}||\boldsymbol{e}_{\boldsymbol{id}}^{a_{src}}||\boldsymbol{v}_{age}]$ is processed through stacked blocks, where age modulates activation via the scaling and shifting: $\hat{\boldsymbol{h}}^{(l+1)} = \boldsymbol{\gamma}^{(l)} \odot \boldsymbol{h}^{(l)} + \boldsymbol{\beta}^{(l)}$, with $\odot$ denoting element-wise multiplication. The modulated features are propagated to produce the raw residual $\boldsymbol{r}_{raw}$. The residual magnitude is further controlled by a learnable age-gap-aware gate: $\psi(\Delta a) = \min(1, \max(0, \theta_1(1 - \exp(-|\Delta a|/\tau_1))))$, where τ_1 is temporal scale and θ_1 is a learnable gain. The final residual is:

$$\boldsymbol{r}^{a_{src}\to a_{trg}} = \psi(\Delta a)\boldsymbol{r}_{raw} \tag{5}$$

with $\psi(\Delta a)$ enforcing smoothly increasing developmental modulation that saturates at large temporal distances while preserving stable gradients for small age differences. Of note, the residual formulation operates at the BOLD signal-dynamics level rather than the FC-edge level, and because FC is computed afterward from generated time series, NDE does not impose monotonic or plateauing FC trajectories.

Optimization via causal distillation. To effectively simulate longitudinal evolution of brain dynamics, NDE is trained under a Teacher-Student distillation framework, where the adapted model $\hat{\boldsymbol{F}}_{\theta^{a_{trg}}}$ (Student) learns to reproduce both BOLD signal predictions and underlying causal dynamics of pre-trained target-age surrogates $\boldsymbol{F}_{\theta^{a_{trg}}}$ (Teacher). The residual transition module thus acts as a causal adaptor, transforming source-age dynamics to target age with behavior and causal distillation.

Given input $\boldsymbol{p}_t^{a_{trg}}$, the behavior distillation encourages the residual-corrected next-

step prediction $\widehat{\boldsymbol{x}}_{t+1}^{a_{trg}}$ to reproduce Teacher's prediction $\widetilde{\boldsymbol{x}}_{t+1}^{a_{trg}}$ at the signal level:

$$\mathcal{L}_{\text{beh}} = \sum_{t=\ell}^{T-1} \left\| \widehat{\boldsymbol{F}}_{\theta^{a_{trg}}}\left(\boldsymbol{p}_t^{a_{trg}}\right) - \boldsymbol{F}_{\theta^{a_{trg}}}\left(\boldsymbol{p}_t^{a_{trg}}\right) \right\|_2^2 \quad (6)$$

To transfer dynamic mechanisms, we distill causal properties at both local and global levels, combining EC matching for local causal structure with Jacobian spectrum distillation for global dynamical stability. Thus, $\mathcal{L}_{\text{cau}} = \mathcal{L}_{\text{EC}} + \mathcal{L}_{\text{dyn}}$,

$$\mathcal{L}_{\text{EC}} = \lambda_{\text{ec}_1}\left(1 - \text{corr}(\text{EC}^{\boldsymbol{F}}, \text{EC}^{\widehat{\boldsymbol{F}}})\right) + \lambda_{\text{ec}_2}\left\|\text{EC}^{\boldsymbol{F}} - \text{EC}^{\widehat{\boldsymbol{F}}}\right\|_2^2, \mathcal{L}_{\text{dyn}} = \left\|\boldsymbol{\sigma}^{\boldsymbol{F}} - \boldsymbol{\sigma}^{\widehat{\boldsymbol{F}}}\right\|_2^2 \quad (7)$$

where $\text{EC}^{\boldsymbol{F}}$ and $\text{EC}^{\widehat{\boldsymbol{F}}}$ denote Teacher and Student EC matrices simulated from Jacobian-based perturbation analysis, $\text{corr}(\cdot)$ is the Pearson correlation over vectorized off-diagonal entries, $\boldsymbol{\sigma}^{\boldsymbol{F}}$ and $\boldsymbol{\sigma}^{\widehat{\boldsymbol{F}}}$ represent leading Jacobian eigen-spectrum components.

To enforce coherent developmental trajectories across ages, we introduce a soft-label contrastive objective operating on free-rollout FC [11], computed from BOLD signals generated autonomously by the adapted surrogate through autoregressive unrolling from random noise, without access to empirical signals. The cosine similarity matrix $\mathcal{M}_{FC}$, computed from subject-wise free-rollout FCs across all age pairs, is aligned with an age-decaying prior similarity matrix $\mathcal{M}_{age} = \left(\exp\left(-(\Delta a)^2/2\tau_{age}^2\right)\right)$, yielding

$$\mathcal{L}_{\text{stru}} = 1/N \sum_{i=1}^{N} \text{KL}(\text{softmax}(\mathcal{M}_{FC}[i]) || \text{softmax}(\mathcal{M}_{age}[i])) \quad (8)$$

Full objective. The final optimization objective is: $\mathcal{L}_{\text{NDE}} = \lambda_{\text{beh}}\mathcal{L}_{\text{beh}} + \lambda_{\text{stru}}\mathcal{L}_{\text{stru}} + \lambda_{\text{cau}}\mathcal{L}_{\text{cau}}$, where λ_{beh} , λ_{stru}, λ_{cau}, λ_{ec_1}, and λ_{ec_2} are trade-off parameters.

2.2 Mechanistic Interrogation of Developmental Dynamics

Training NDE on a typically developing (TD) cohort establishes a normative generative model of longitudinal brain dynamics with causal developmental inference capability, uniquely positioning NDE as a framework for identifying atypical maturation and probing underlying developmental mechanisms.

Trajectory divergence identification. For at-risk populations (e.g. infants with prenatal drug exposure (PDE)), paired longitudinal scans spanning a younger source age and a later target age are evaluated using the TD-trained NDE. Initialized with source-age data, NDE simulates normative developmental trajectories toward the target age. Divergence is quantified by comparing simulated FC with observations, where significant discrepancies reveal connections (edges) deviating from typical maturation (**Fig. 2a**).

Counterfactual developmental knobs. To mechanistically interrogate how early localized perturbations reshape FC over development, we introduce counterfactual developmental knobs (i.e., brain region-level FC drivers) for transition window $a_{src} \rightarrow a_{trg}$. This design modulates the ***future*** developmental trajectory with ***current*** spatial perturbations. As in **Fig. 2b**, with NDE parameters fixed, a non-optimized intervention vector $\boldsymbol{g} \in \mathbb{R}^K$ is introduced as a differentiable perturbation variable to modulate region-specific residual updates originating from the source age:

$$\widehat{\boldsymbol{x}}_{t+1}^{a_{trg}} = \boldsymbol{F}_{\theta^{a_{src}}}\left(\boldsymbol{z}_t\right) + \boldsymbol{g} \odot \boldsymbol{r}^{a_{src} \rightarrow a_{trg}}, \boldsymbol{z}_t \in \mathcal{N}(0, e\mathbf{I}), e \in (0,1) \quad (9)$$

Free rollout of the perturbed system from random noise $\boldsymbol{z}_t$ produces counterfactual

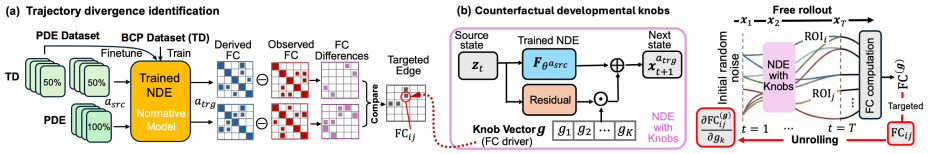

Fig. 2. (a) Trajectory divergence identification and (b) Counterfactual developmental knobs.

BOLD signal, from which the intervened target-age $\mathrm{FC}^{(\boldsymbol{g})}$ is computed. To quantify how an early perturbation at region k influences a future edge FC_{ij}, we define a leverage score as the local sensitivity at the neutral setting of $\boldsymbol{g} = \mathbf{1}$: $\mathrm{Imp}_k(\mathrm{FC}_{ij}) = \partial \mathrm{FC}_{ij}^{(\boldsymbol{g})} / \partial g_k \,|_{\boldsymbol{g}=\mathbf{1}}$. Computed efficiently via automatic differentiation through the unrolled dynamics, ranking $|\mathrm{Imp}_k|$ identifies dominant developmental knobs while the sign of Imp_k determines modulation direction. A **positive FC driver** ($\mathrm{Imp}_k > 0$) amplifies the FC_{ij} at the target age when region k is upregulated at source age, whereas a **negative FC driver** ($\mathrm{Imp}_k < 0$) weakens it, resembling excitation-inhibition-like modulation of developmental connectivity. This formulation treats the residual pathway as a controllable causal adaptor, enabling identification of brain regions whose early modulation reshapes FC outcomes across development.

3 Experiments

3.1 Data Description

Experimental validation was performed on two longitudinal rs-fMRI cohorts spanning early neurodevelopment from birth to 24 months. The primary dataset was from the publicly available UNC/UMN Baby Connectome Project (BCP) [15], comprising 1,151 longitudinal scans from 213 typically developing (TD) infants, enabling robust modeling of normative developmental dynamics. To assess translational and clinical utility, the framework was further evaluated on an independent in-house cohort (PDE dataset) containing 449 scans from 131 infants, including 96 infants with prenatal drug exposure (PDE) and 35 TD controls. This dual-cohort design enables evaluation across both normative and at-risk developmental trajectories. All rs-fMRI data underwent rigorous preprocessing and quality control using state-of-the-art infant-specific pipelines [16-19].

3.2 Validation of NDE

As no existing methods explicitly model longitudinal transitions between neural surrogates and conventional FC prediction does not model causal dynamics, evaluation was conducted through internal benchmarking and ablation studies. The surrogate ANN in our NDE was implemented as a multi-layer perceptron (MLP). It was compared against the following baselines and architectural variants: (1) Baseline: source-age surrogate ($\boldsymbol{F}_{\theta^{a_{src}}}$) without developmental residual modeling. (2) NDE-RNN: recurrent neural network surrogate ANN backbone. (3) NDE-CNN: convolutional neural network [11] surrogate ANN backbone. To assess the contribution of key optimization components, the full NDE was further compared with three ablated variants: (4) NDE w/o C: without

soft-label contrastive loss. (5) NDE w/o b: without behavior distillation loss. (6) NDE-FC: direct FC optimization using L_2 and correlation losses without causal distillation.

Model performance is evaluated by **1) Predictive accuracy**: FC, as a key developmental neuroimaging biomarker, serves as the primary evaluation target. Performance is quantified using the Pearson Correlation Coefficient (PCC) between target empirical FC and NDE-simulated FC under teacher-forced prediction (FC_{tf}, from $\boldsymbol{p}_t^{a_{trg}}$) and free rollouts ($\mathrm{FC}_{free}$, from noise $\boldsymbol{z}_t$) [11]. EC preservation is assessed via PCC between teacher and student-derived EC matrices. **2) Developmental fidelity**. Assessed through age prediction and functional fingerprinting [20] based on simulated FC and EC, measuring preservation of developmental age effect and subject identity. Developmental specificity is further evaluated by computing PCC between the NDE-simulated FC_{free} for the same subject at different target ages to verify age-dependent differentiation.

Experimental settings. All models were implemented in PyTorch and optimized using AdamW with subject-wise batching. The identity encoder features a 3-layer GCN (hidden/output 256-d, dropout 0.1) with normalized message passing and global mean pooling, while the residual dynamics network employs a two-block FiLM architecture (512-d, SiLU activation). Learning rates were managed via a ReduceLROnPlateau scheduler, initialized at 5×10^{-4} for the residual network and 5×10^{-5} for the identity encoder (weight decay 10^{-4}). Loss weights were set to λ_{beh}=3, λ_{ec_1}= λ_{ec_2}=3.5, λ_{cau} = 0.1, and λ_{stru} = 0.1. Free-rollout FC was computed from 200-step autoregressive rollouts initialized from Gaussian noise and averaged across four random seeds, while Jacobian spectrum regularization matched the leading singular value. For PDE dataset fine-tuning, the identity encoder was frozen, and the learning rate reduced to 1×10^{-5}.

Table 1. Comparison of NDE with six competing methods on BCP dataset. (Mean ± std. reported for Predictive Accuracy; Mean for Developmental Fidelity; ↓ indicates lower is better)

	Predictive Accuracy (PCC)				**Developmental Fidelity**			
					Age prediction		Fingerprinting	
	FC_{free}	FC_{tf}	EC	Specificity↓	FC_{free}	EC	FC_{free}	EC
Baseline	0.645 ± 0.012	0.765 ± 0.015	0.232 ± 0.049	1.000 ± 0.000	0.552	0.351	-	-
NDE-RNN	0.650 ± 0.015	0.792 ± 0.013	0.363 ± 0.021	0.662 ± 0.014	**0.635**	0.721	0.156	0.102
NDE-CNN	0.501 ± 0.034	0.592 ± 0.021	0.298 ± 0.025	0.719 ± 0.018	0.466	-0.030	0.454	0.452
NDE w/o C	**0.693 ± 0.015**	0.830 ± 0.016	0.489 ± 0.015	0.752 ± 0.025	0.607	0.716	0.605	0.461
NDE w/o b	0.614 ± 0.024	0.802 ± 0.009	0.490 ± 0.012	0.659 ± 0.012	0.521	0.712	0.452	**0.515**
NDE-FC	0.660 ± 0.017	0.795 ± 0.013	0.373 ± 0.013	0.682 ± 0.009	0.494	0.390	0.609	0.473
Our NDE	0.681 ± 0.013	**0.839 ± 0.011**	**0.502 ± 0.015**	**0.643 ± 0.010**	0.621	**0.738**	**0.615**	0.507

Results. Rs-fMRI signals were extracted using the Schaefer-100 functional parcellation [21] and evaluated with subject-wise five-fold cross-validation. On the BCP dataset (**Table 1**), NDE achieved the best overall performance, with paired permutation tests confirming significant gains over competing methods across evaluation tasks (p < 0.001). FC prediction improved from 0.765/0.645 (FC_{tf} / FC_{free} , baseline) to 0.839/0.681, while EC alignment increased from 0.232 to 0.502, indicating effective prediction of causal dynamics. NDE also achieved the highest age prediction accuracy

(0.621 using FC_{free}, 0.738 using EC), and functional fingerprinting further improved to 0.615 (FC_{free}) and 0.507 (EC), confirming preservation of developmental age effects and individualized neural signatures. Ablation studies showed performance degradation with key components removed. Direct FC optimization (NDE-FC) notably reduced EC alignment (0.373) and age prediction (0.390), underscoring the importance of causal distillation. Removing contrastive or behavioral supervision also degraded performance. Notably, although NDE w/o C slightly improved FC_{free}, its developmental specificity severely degraded, indicating compromised developmental differentiation.

Moreover, the BCP-trained model, fine-tuned on 68 longitudinal scans from 18 TD infants (50% of PDE TD controls, **Fig. 2a**), showed reduced performance in the PDE cohort (**Fig. 3a**), suggesting divergence from typical maturation in infants with prenatal drug exposure. Divergent FC was predominantly localized within the Default Mode (Def) and Control (Con) networks (**Fig. 3b**), suggesting disrupted maturation of higher-order regulatory systems in the PDE cohort, consistent with prior neurodevelopmental findings [22, 23]. Counterfactual developmental analysis further identified candidate FC drivers (**Fig. 3c-e**) whose modulation could steer abnormal connectivity toward normative trajectories, revealing shared motor-visual regions as key contributors to developmental modulation. Importantly, dominant FC drivers differed between the 3→6 months (**Fig. 3c**) and 3→9 months (**Fig. 3d**) transition window, indicating that causal control of the network is temporally reconfigured rather than static. This transition-specific redistribution of FC drivers demonstrates that effective intervention is inherently stage-dependent and should align with the dynamically evolving network architecture. Moreover, substantial inter-individual heterogeneity was observed in the spatial distribution of FC drivers (**Fig. 3e**) even within the same developmental window, highlighting the necessity of individualized intervention strategies. Overall, these results establish that NDE advances beyond predictive modeling by enabling mechanistic characterization of developmental divergence and corresponding modulations.

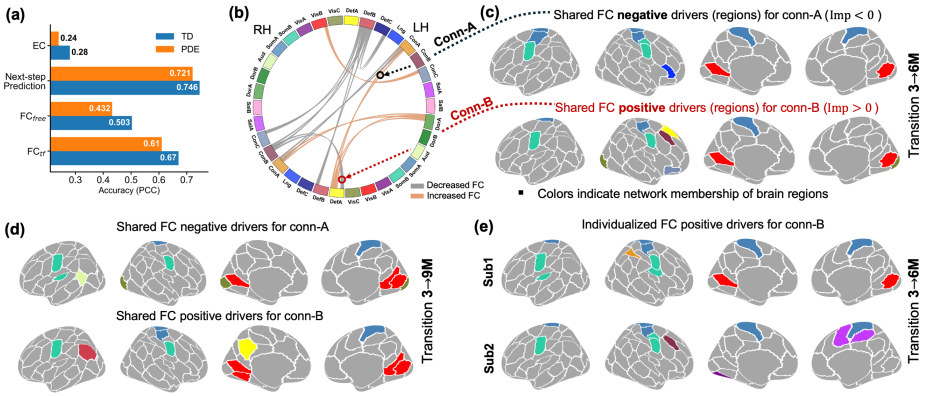

Fig. 3. Validation of NDE on the PDE dataset. (a) Fine-tuning performance; (b) Top 20 divergent connections in PDE cohort; (c) Top 10 shared FC drivers with the highest leverage scores across PDE infants over the 3→6M (months) developmental transition; (d) Top 10 shared FC drivers across PDE infants (3→9M); (e) Top 10 FC drivers for two exemplar infants with PDE (3→6M).

4 Conclusion

We present the Neuro-Developing Engine (NDE), the first surrogate longitudinal framework that moves beyond conventional static connectivity prediction toward modeling individualized brain development as a dynamic and causally interpretable process. By enabling counterfactual interrogation of developmental trajectories, NDE supports developmental causal inference and reveals network mechanisms underlying both typical and atypical maturation. Several limitations should be noted. NDE requires scan-specific surrogate estimation, and although these lightweight surrogates are independently trainable, their stability under limited or noisy rs-fMRI time points warrants further evaluation. PDE-domain adaptation relied on a limited number of TD infants, so the resulting neuroscience findings should be interpreted cautiously despite protocol- and age-matched scans and conservative fine-tuning. In addition, the age-gap gate improves training stability but may reduce sensitivity to abrupt dynamical shifts. Future work will validate NDE in larger cohorts and broader developmental, aging, and clinical settings. Overall, although developed for the rapid nonlinear changes of early infancy, this causal distillation paradigm provides a generalizable foundation for mechanistic modeling of human brain development and aging across the lifespan.

Acknowledgments. This work was supported in part by NIH grants (AG075582, NS128534, EB037388, and NS135574).

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Wen, X., Wang, R., Yin, W., Lin, W., Zhang, H., Shen, D.: Development of dynamic functional architecture during early infancy. Cerebral Cortex 30, 5626-5638 (2020)
2. Yin, W., Li, T., Wu, Z., Hung, S.-C., Hu, D., Gui, Y., Cho, S., Sun, Y., Woodburn, M.A., Wang, L.: Charting brain functional development from birth to 6 years of age. Nature Human Behaviour 9, 1246-1259 (2025)
3. Yin, W., Chen, M.-H., Hung, S.-C., Baluyot, K.R., Li, T., Lin, W.: Brain functional development separates into three distinct time periods in the first two years of life. NeuroImage 189, 715-726 (2019)
4. Parkes, L., Satterthwaite, T.D., Bassett, D.S.: Towards precise resting-state fMRI biomarkers in psychiatry: synthesizing developments in transdiagnostic research, dimensional models of psychopathology, and normative neurodevelopment. Current Opinion in Neurobiology 65, 120-128 (2020)
5. Tapia Medina, M.G., Cosío-Guirado, R., Peró-Cebollero, M., Cañete-Massé, C., Villuendas-González, E.R., Guàrdia-Olmos, J.: The clinical relevance of healthy neurodevelopmental connectivity in childhood and adolescence: a meta-analysis of resting-state fMRI. Frontiers in neuroscience 19, 1576932 (2025)
6. Hiremath, C.S., Sagar, K.J.V., Yamini, B., Girimaji, A.S., Kumar, R., Sravanti, S.L., Padmanabha, H., Vykunta Raju, K., Kishore, M.T., Jacob, P.: Emerging behavioral and neuroimaging biomarkers for early and accurate characterization of autism spectrum disorders: a systematic review. Translational psychiatry 11, 42 (2021)
7. Yu, X., Hu, D., Zhang, L., Huang, Y., Wu, Z., Liu, T., Wang, L., Lin, W., Zhu, D., Li, G.: Longitudinal infant functional connectivity prediction via conditional intensive triplet network. In: International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 255-264. Springer Nature Switzerland Cham, (2022)
8. Bessadok, A., Nebli, A., Mahjoub, M.A., Li, G., Lin, W., Shen, D., Rekik, I.: A few-shot learning graph multi-trajectory evolution network for forecasting multimodal baby connectivity development from a baseline timepoint. In: International Workshop on PRedictive Intelligence In MEdicine, pp. 11-24. Springer, (2021)
9. Ciric, R., Rosen, A.F., Erus, G., Cieslak, M., Adebimpe, A., Cook, P.A., Bassett, D.S., Davatzikos, C., Wolf, D.H., Satterthwaite, T.D.: Mitigating head motion artifact in functional connectivity MRI. Nature protocols 13, 2801-2826 (2018)
10. Birn, R.M., Cornejo, M.D., Molloy, E.K., Patriat, R., Meier, T.B., Kirk, G.R., Nair, V.A., Meyerand, M.E., Prabhakaran, V.: The influence of physiological noise correction on test–retest reliability of resting-state functional connectivity. Brain connectivity 4, 511-522 (2014)
11. Luo, Z., Peng, K., Liang, Z., Cai, S., Xu, C., Li, D., Hu, Y., Zhou, C., Liu, Q.: Mapping effective connectivity by virtually perturbing a surrogate brain. Nature Methods 22, 1376-1385 (2025)
12. Zhang, Y., Liu, D., Liang, Z., Cheng, J., Lou, K., Duan, J., Gao, T., Hu, B., Liu, Q.: Artificial intelligence as a surrogate brain: bridging neural dynamical models and data. National Science Review 13, nwaf457 (2026)
13. Schirner, M., Domide, L., Perdikis, D., Triebkorn, P., Stefanovski, L., Pai, R., Popa, P., Valean, B., Palmer, J., Langford, C.: Brain modelling as a service: the virtual brain on EBRAINS. arXiv preprint arXiv:2102.05888 (2021)
14. Perez, E., Strub, F., De Vries, H., Dumoulin, V., Courville, A.: Film: Visual reasoning with a general conditioning layer. In: Proceedings of the AAAI conference on artificial intelligence. (2018)

15. Howell, B.R., Styner, M.A., Gao, W., Yap, P.-T., Wang, L., Baluyot, K., Yacoub, E., Chen, G., Potts, T., Salzwedel, A.: The UNC/UMN Baby Connectome Project (BCP): An overview of the study design and protocol development. NeuroImage 185, 891-905 (2019)
16. Wang, L., Wu, Z., Chen, L., Sun, Y., Lin, W., Li, G.: iBEAT V2. 0: a multisite-applicable, deep learning-based pipeline for infant cerebral cortical surface reconstruction. Nature protocols 18, 1488-1509 (2023)
17. Lim, M., Heo, K.-S., Kim, J.-M., Kang, B., Lin, W., Zhang, H., Shen, D., Kam, T.-E.: A unified multi-modality fusion framework for deep spatio-spectral-temporal feature learning in resting-state fMRI denoising. IEEE Journal of Biomedical and Health Informatics 28, 2067-2078 (2024)
18. Smith, S.M., Jenkinson, M., Woolrich, M.W., Beckmann, C.F., Behrens, T.E., Johansen-Berg, H., Bannister, P.R., De Luca, M., Drobnjak, I., Flitney, D.E.: Advances in functional and structural MR image analysis and implementation as FSL. Neuroimage 23, S208-S219 (2004)
19. Avants, B.B., Tustison, N.J., Song, G., Cook, P.A., Klein, A., Gee, J.C.: A reproducible evaluation of ANTs similarity metric performance in brain image registration. Neuroimage 54, 2033-2044 (2011)
20. Hu, D., Wang, F., Zhang, H., Wu, Z., Zhou, Z., Li, G., Wang, L., Lin, W., Li, G., Consortium, U.U.B.C.P.: Existence of functional connectome fingerprint during infancy and its stability over months. Journal of Neuroscience 42, 377-389 (2022)
21. Schaefer, A., Kong, R., Gordon, E.M., Laumann, T.O., Zuo, X.-N., Holmes, A.J., Eickhoff, S.B., Yeo, B.T.: Local-global parcellation of the human cerebral cortex from intrinsic functional connectivity MRI. Cerebral cortex 28, 3095-3114 (2018)
22. Jiang, W., Merhar, S.L., Zeng, Z., Zhu, Z., Yin, W., Zhou, Z., Wang, L., He, L., Vannest, J., Lin, W.: Neural alterations in opioid-exposed infants revealed by edge-centric brain functional networks. Brain Communications 4, fcac112 (2022)
23. De Asis-Cruz, J., Kim, J.-H., Kapse, K., Wu, Y., Merhar, S.L., Bann, C.M., Newman, J.E., Mack, N., DeMauro, S.B., Ambalavanan, N.: Disrupted brain connectivity in newborns following antenatal opioid exposure. Biological Psychiatry: Cognitive Neuroscience and Neuroimaging (2025)