

Physics-Informed AI Framework Enables Generalizable and Reliable Cerebrovascular Hemodynamic Profiling

Fen Miao

University of Electronic Science and Technology of China

Chen Chen

Hong Kong Polytechnic University

Yuan Lin

University of Electronic Science and Technology of China

Yun Zhang

Sun Yat-sen University

Jiannong Cao

Hong Kong Polytechnic University

Yingjie He

University of Electronic Science and Technology of China

Jiaqi Huang

University of Electronic Science and Technology of China

Honglei Zhao

Fuwai Hospital, Chinese Academy of Medical Sciences·Shenzhen

Jun Yang

University of Electronic Science and Technology of China

Huiying Liang

Fudan University

Ye Li

`ye.li@siat.ac.cn`

Shenzhen Institutes of Advanced Technology

Article

Keywords:

Posted Date: July 16th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-9990730/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.
[Read Full License](#)

Additional Declarations: No competing interests reported.

Physics-Informed AI Framework Enables Generalizable and Reliable Cerebrovascular Hemodynamic Profiling

Fen Miao^{1†}, Chen Chen^{2,3†}, Yuan Lin^{1†}, Yun Zhang⁴, Jiannong Cao³, Yingjie He¹, Jiaqi Huang¹, Honglei Zhao⁵, Jun Yang^{1*}, Huiying Liang^{6*}, Ye Li^{2,7*}

¹Shenzhen Institute for Advanced Study, University of Electronic Science and Technology of China, Shenzhen, China

²Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China

³Hong Kong Polytechnic University, Hong Kong, China

⁴School of Electronics and Communication Engineering, Sun Yat-sen University, Shenzhen, China

⁵Fuwai Hospital, Chinese Academy of Medical Sciences·Shenzhen, Shenzhen, China

⁶Intelligent Medicine Institute, Fudan University, Shanghai, China

⁷Shenzhen University of Advanced Technology, Shenzhen, China.

*Corresponding author. Email: junyang@uestc.edu.cn; lianghuiying@fudan.edu.cn; ye.li@siat.ac.cn;

[†] These authors contributed equally to this work.

Abstract:

Accurate spatio-temporal profiling of cerebrovascular hemodynamics is essential for neurovascular disease management. However, clinical translation of artificial intelligence in this domain is severely bottlenecked by the "sim-to-real" gap: models trained on idealized *in silico* data struggle to adapt to the highly variable and complex vascular anatomies of real-world patients. To bridge this gap, we present 4DHemoX, a reliable physics-informed AI framework designed for generalizable hemodynamic profiling. This framework is powered by 4DHemoDB, a comprehensive hybrid dataset comprising 168,000 high-fidelity *in silico* frames across 280 patient-derived vascular geometries, alongside 800 *in vivo* 4D-flow MRI frames. At its core is 4DHemoFormer, a specialized neural architecture that embeds Navier-Stokes equations directly into its learning process, enhancing the model's generalizability to complex flow regimes. 4DHemoFormer demonstrates robust generalization across unseen vascular geometries (nRMSE: 2.39×10^{-2}) and successfully overcomes the sim-to-real barrier on clinical data. It yields a strong zero-shot baseline (nRMSE: 6.28×10^{-2}) and exhibits remarkable data efficiency, requiring only 10% of *in vivo* data to calibrate to clinical MRI measurements (nRMSE: 3.51×10^{-2}). Ultimately, 4DHemoX provides a reliable platform for real-world hemodynamic profiling, paving the way for non-invasive biomarker quantification and proactive neurovascular interventions.

INTRODUCTION

Accurate spatio-temporal profiling of cerebrovascular hemodynamics is essential for elucidating the complex biomechanical forces, such as velocity and pressure fluctuations, that drive cerebrovascular anomalies like aneurysmal rupture and ischemic stroke¹⁻³. Traditionally, clinical workflows rely on static morphological assessments, which inherently fail to capture the complex, patient-specific fluid dynamics triggering these acute events. Spatio-temporal hemodynamic profiling bridges this critical gap by unlocking high-resolution biomechanical biomarkers that remain invisible to standard clinical imaging. Consequently, this capability drives a paradigm shift in clinical practice: transitioning from reactive anatomical diagnostics to proactive, functional forecasting for personalized neurovascular care^{4,5,6}.

Currently, acquiring spatio-temporal hemodynamics relies on two primary modalities, both of which face significant translational barriers. Four-dimensional (4D) flow magnetic

resonance imaging (MRI) provides a non-invasive means to capture time-resolved *in vivo* hemodynamics^{7,8}. However, its widespread utility is often hindered by high-costs, prolonged acquisition times, limited spatio-temporal resolution, and susceptibility to motion artifacts⁹. Conversely, Computational Fluid Dynamics (CFD) serves as a robust *in silico* alternative¹⁰⁻¹³. By numerically solving the Navier-Stokes (NS) equations using discretization techniques such as finite volume¹⁴ or finite element¹⁵ methods, CFD reconstructs full, time-varying velocity and pressure fields with unprecedented spatio-temporal resolution. Yet, despite its high fidelity, applying CFD to intricate, patient-specific cerebrovascular structures is computationally prohibitive, presenting a formidable bottleneck to routine clinical integration^{16,17}.

Machine learning has emerged as a transformative approach to bypass CFD's computational bottleneck, enabling rapid inference of high-dimensional flow fields¹⁸⁻²⁰. While foundational AI frameworks, such as physics-informed neural networks (PINNs)²¹⁻³¹ and neural operators (e.g., FNO)³²⁻⁴⁰, provide a robust mathematical basis for fluid modeling, current approaches remain trapped in a severe "sim-to-real" gap. Predominant architectures are largely confined to idealized, simulated 2D geometries, inherently failing to capture the complex irregularities of patient-specific cerebrovascular anatomies and the nuances of real-world clinical data. Furthermore, while Transformer architectures offer a promising solution due to their inherent adaptability to irregular structures⁴¹⁻⁴⁴, their potential for reconstructing complex 3D spatio-temporal hemodynamics remains largely untapped^{45,46}. This architectural bottleneck is further compounded by a critical lack of comprehensive benchmarks; although preliminary baselines like CFDBench⁴⁷ and FlowBench⁴⁸ exist, the field is paralyzed by severe data scarcity and an evaluation gap that obscures model generalization across patient-derived anatomies and real-world MRI data^{49,50}. Collectively, these deficiencies frequently result in weak baselines and obscure true model generalization⁵⁰. This underscores an urgent need for an integrated framework that encompasses patient-specific databases, advanced AI models, and unified validation protocols to bridge the gap between *in silico* development and clinical deployment.

Here, we introduce 4DHemoX, an integrated, physics-informed generalizable framework designed to facilitate the clinical translation of cerebrovascular hemodynamic modeling. By unlocking high-resolution spatio-temporal velocity and pressure fields, 4DHemoX provides

critical biomechanical biomarkers to enable preemptive interventions and redefine personalized neurovascular care (Fig. 1 a). This work makes three primary contributions:

(1) To explicitly bridge the "sim-to-real" gap, we present a cross-domain database 4DHemoDB. 4DHemoDB integrates high-fidelity CFD simulations, encompassing diverse, patient-specific cerebrovascular structures (including the common carotid arteries [CCA], external carotid arteries [ECA], and internal carotid arteries [ICA], and vertebral artery [VA]) under varying physiological boundary conditions, with *in vivo* 4D-flow MRI acquisitions. Beyond spatio-temporal forecasting, this database serves as a foundational benchmark for broader translational applications such as hemodynamic super-resolution.

(2) To process the highly unstructured spatial graphs inherent to complex vasculature, we propose 4DHemoFormer, a physics-informed generalizable Transformer architecture. By encoding physical constraints directly into its attention mechanism and loss landscape, this tailored architecture enables robust generalization across highly heterogeneous data and irregular anatomical structures.

(3) We formulate a multi-tiered evaluation paradigm to rigorously assess AI-based fluid dynamics models across three challenging dimensions critical for clinical deployment: temporal extrapolation (enabling extended hemodynamic forecasting beyond the training horizon), cross-geometry generalization (ensuring adaptability to unseen vascular geometries), and cross-domain transferability (bridging the sim-to-real gap for application to real-world clinical MRI data). Leveraging 4DHemoDB, we comprehensively benchmark 4DHemoFormer against state-of-the-art baselines (including PINNs²¹, FNO³³, and DeepONet³⁴), providing critical insights into performance trade-offs and the translational viability of AI models for clinical use.

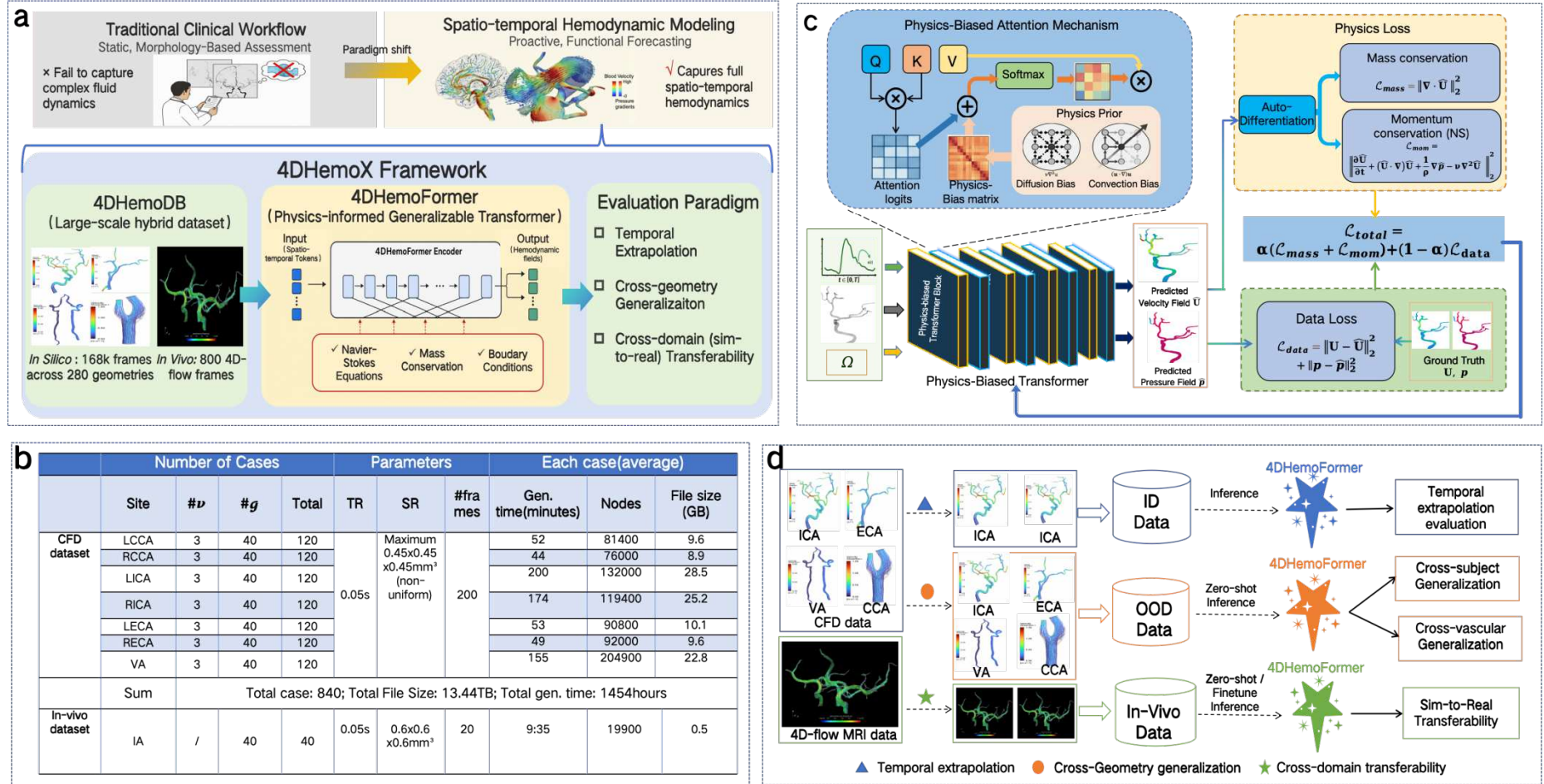


Fig. 1. | Illustrations of the 4DHemoX. a, Overview of the motivation and 4DHemoX framework. b, Data description of the 4DHemoDB. c, Illustration of the overall architecture of the proposed 4DHemoFormer. d, Comprehensive evaluation protocol. 4DHemoFormer is evaluated on in-distribution (ID), out-of-distribution (OOD) and real-world settings to evaluate the temporal extrapolation ability (enabling extended hemodynamic forecasting beyond the training horizon), cross-geometry generalization (ensuring adaptability to unseen vascular geometries), and cross-domain transferability (bridging the sim-to-real gap for application to real-world clinical MRI data). LCCA, left Common Carotid Artery; RCCA, right Common Carotid Artery; LICA, left Internal Carotid Artery; RICA, right Internal Carotid Artery; LECA, left External Carotid Artery; RECA, right External Carotid Artery; VA, Vertebral Artery; TR, temporal resolution; SR, spatial resolution. IA, Intracranial artery.

RESULTS

4DHemoDB, a large-scale cross-domain cerebrovascular hemodynamics database

4DHemoDB is a large-scale, hybrid *in silico* - *in vivo* cerebrovascular hemodynamics database, established as a foundational resource for developing and rigorously evaluating AI-based models in spatio-temporal fluid forecasting. To construct 4DHemoDB, we enrolled a cohort of 40 subjects (median age, 49 years; interquartile range [IQR], 32.5 – 58), including 7 patients presenting with carotid artery stenosis [detailed demographics are summarized in Supplementary Table 1]. Patient-specific cerebrovascular geometries, 4D-flow MRI data, and velocity waveforms from the bilateral common and internal carotid arteries, were acquired for each subject (detailed collection protocols are detailed in the Methods). The database comprises two complementary partitions: (1) a Computational Hemodynamics dataset (CFD dataset), generated via high-fidelity numerical simulations across diverse patient-specific cerebrovascular structures (including CCA, ICA, ECA and VA) with varied boundary conditions and blood viscosity (Extended data Fig. 1a); and (2) an *in vivo* Hemodynamics dataset (In-Vivo dataset), captured directly from intracranial arteries using 4D-flow MRI sequences (Extended Data Fig. 1b). The final 4DHemoDB with its key statistics are summarized in Fig. 1b. 4DHemoDB offers substantial advantages over existing benchmarks in terms of patient-specific geometric complexity, and the inclusion of real-world clinical measurements, establishing it as a comprehensive and physiologically relevant benchmark for hemodynamic analysis (see details in extended data Table 1).

CFD Dataset. Hemodynamics is fundamentally governed by the Navier-Stokes (NS) equations⁵¹, which enforce mass and momentum conservation. In this study, we employ numerical methods to solve these PDEs, simulating the complex interplay of blood velocity, pressure, and viscosity to generate our high-fidelity CFD dataset. Blood was modeled as an incompressible Newtonian fluid, with the equations written as:

$$\begin{cases} \frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla) \mathbf{u} = -\frac{1}{\rho} \nabla p + \nu \nabla^2 \mathbf{u} \\ \nabla \cdot \mathbf{u} = 0 \end{cases} \quad (1)$$

where ρ is the density and ν is the blood viscosity, $\mathbf{u} = (u, v, w)^T$ is the velocity field, and p is the pressure field. We considered a constant density of 1060 kg/m^3 and a dynamic viscosity of $\nu = \{0.00277, 0.00345, 0.00413 \text{ Pa.s}\}$, consistent with standard practices in cerebrovascular hemodynamics^{52,53}. The details for generating CFD dataset are presented in Methods section. The CFD dataset contains a total of 168,000 temporal frames, derived from 840 simulation cases. These simulations are performed across 280 unique patient-specific vascular structures at 7 sites (left/right CCA, left/right ICA, left/right ECA and VA), incorporating 3 different blood viscosity values. Each simulation has a duration of 10 seconds, with data sampled at a temporal resolution of 0.05s. The mesh size is non-uniform, with the maximum element edge length not exceeding 0.45 mm.

In-Vivo Dataset. *In vivo* 4D-flow MRI data are collected by a non-invasive imaging technique that captures time-resolved, three-dimensional blood flow velocities within the vasculature. A phase-contrast MRI sequence with three-directional velocity encoding is employed (see details in Method section). The In-Vivo dataset includes 800 temporal frames collected from 40 subjects, each corresponding to a unique intracranial artery. Each acquisition captures a full cardiac cycle that divided into 20 frames, providing real-world hemodynamic profiles. The voxel size is $0.6 \times 0.6 \times 0.6 \text{ mm}^3$.

4DHemoFormer design and evaluation settings

In the following, we illustrate the problem set-up and the architecture of 4DHemoFormer, which we further extend to large-scale validation.

Problem set-up. A solution to the fluid dynamic problem is a vector-valued function $\mathbf{U}: \mathcal{G} \times \mathcal{T} \times \Omega \rightarrow \mathbb{R}^{d_u}$, defined over the spatial domain $\mathcal{G} \subset \mathbb{R}^d$, time domain $\mathcal{T}$, and parameter space Ω (such as the blood viscosity, Reynolds number, or forcing terms). For learning spatio-temporal hemodynamics, $\mathbf{U} \in \mathbb{R}^4$ represent the state including x -direction velocity $u \in \mathbb{R}$, y -direction velocity $v \in \mathbb{R}$, z -direction velocity $w \in \mathbb{R}$, and pressure $p \in \mathbb{R}$ at a given location $g = (x, y, z) \in \mathcal{G}$ and moment $t \in \mathcal{T}$, subject to the governing

equations and parameter space $\Omega = (\rho, \nu, \mathcal{B})$. Where ρ is the fluid density, μ is the fluid viscosity, $\mathcal{B}$ is the inflow condition.

Our goal is to train a physics-aware surrogate model that predicts $\mathbf{U}(\cdot, t)$ from initial conditions $\mathbf{U}(\cdot, 0)$, and parameter space Ω while ensuring consistency with physical laws. We focus on two tasks: single-step prediction and multi-step prediction.

For Single-step prediction task, the model f_θ is trained to predict:

$$\hat{\mathbf{U}}(\cdot, t) = f_\theta(\mathbf{U}(\cdot, t-1), \Omega) \quad (2)$$

For multi-step prediction task, the model directly outputs the solution for a query location and time step t :

$$\hat{\mathbf{U}}(\cdot, t) = f_\theta(\mathbf{U}(\cdot, 0), \Omega) \quad (3)$$

In learning, we take a dataset $\mathcal{D}$ comprising of discretized solutions $\mathcal{D} := \{\mathbf{U}(\cdot, [0:t_{\max}])\}$ to find θ^* that minimizes the loss function $\mathcal{L}$.

$$\theta^* = \arg \min_{\theta} \sum_{t=1}^{t_{\max}} \mathcal{L}(\mathbf{U}(\cdot, t), \hat{\mathbf{U}}(\cdot, t)) \quad (4)$$

The 4DHemoFormer framework. We propose 4DHemoFormer, a physics-informed framework designed to predict spatio-temporal velocity and pressure fields within complex vascular structures. 4DHemoFormer integrates Navier-Stokes constraints directly into the attention mechanism and loss landscape of the transformer architecture. As presented in Fig. 1c, the architecture is composed of a Physics-Biased Transformer-based encoder for state estimation and a Physics-Informed Regularization module that enforces consistency with the underlying governing equations. In physics-biased transformer encoder module, a Physics-Bias matrix is explicitly injected into the attention logits. This bias matrix acts as a strong architectural inductive bias, derived from the physical priors of the governing fluid equations. Another distinguishing feature of this architecture is the dual-pathway optimization strategy, which constrains the neural network output using both observational data and the governing physical laws. The network is trained using a composite loss function $\mathcal{L}_{total}$ that balances data

fidelity with physical laws:

$$\mathcal{L}_{total} = (1 - \alpha) \mathcal{L}_{data} + \alpha \mathcal{L}_{physics} \quad (5)$$

where the data Loss ($\mathcal{L}_{data}$) is a weighted mean square error loss comparing predicted fields to ground truth. $\mathcal{L}_{physics}$ is the Navier-Stokes Residual Loss penalizes deviations from the conservation of momentum and conservation of mass. The details of the methodology and implementation can be found in Methods.

Evaluation settings. To comprehensively assess the performance of 4DHemoFormer, we designed a systematic evaluation scheme comprising three distinct settings: temporal extrapolation (extended hemodynamic forecasting beyond the training horizon), cross-geometry generalization (ensuring adaptability to unseen vascular geometries), and cross-domain transferability (bridging the sim-to-real gap for application to real-world clinical MRI data). (Fig. 1d and Extended Data Fig. 2; see Methods for details). In temporal extrapolation setting, the model is trained on initial time frames and evaluated on subsequent, unseen frames for each vascular site. For the cross-geometry setting, we consider two scenarios to simulate realistic clinical challenges: (1) cross-subject generalization, where the model is evaluated on unseen individuals to address inter-subject variability; and (2) cross-vascular generalization, where the model is trained on three vascular sites and evaluated on a fourth unseen site to account for anatomical discrepancies across different vascular beds. Finally, to demonstrate clinical applicability via cross-domain transfer, we evaluated the CFD-trained model on the In-Vivo dataset, rigorously testing its sim-to-real generalization capabilities. For comparison, we included 3 baseline architectures: PINN²¹, DeepONet³³ and FNO³⁴ (see more details in Methods). We employ the normalized relative mean square error (nRMSE) and correlation coefficient (CC) as the main evaluation metrics (see the definitions in Methods).

Temporal extrapolation for extended hemodynamic forecasting

To evaluate 4DHemoFormer’s performance on forecasting extended hemodynamics, we conduct the temporal extrapolation evaluation for both single-step prediction and multi-step prediction task across the CFD dataset of 4DHemoDB. Fig. 2a indicates that data from the training subset and test subset have the similar distribution.

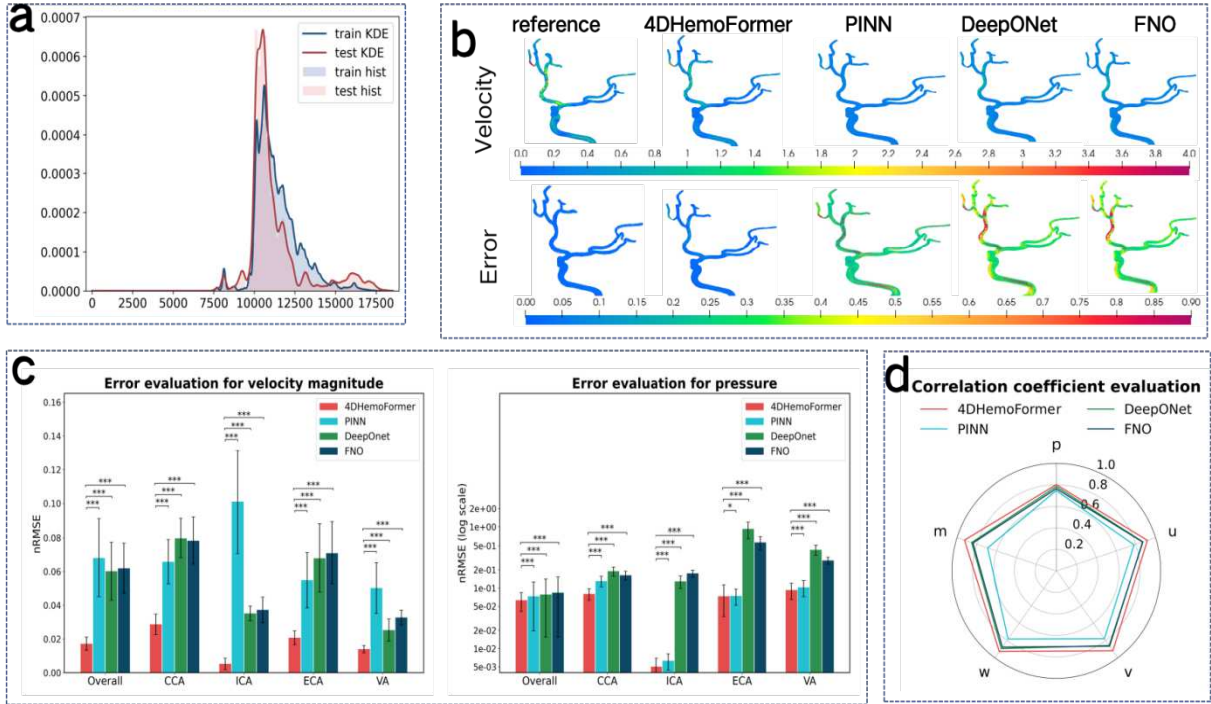


Fig. 2. | Performance of temporal extrapolation in single-step prediction task. a. Data distribution in the training and testing dataset, with the pressure component as the example. b. Comparison of the predicted and reference hemodynamic velocity fields, using the ICA as a representative example. From left to right, the panels represent the numerical simulation (denoted as Ref.), the 4DHemoFormer prediction, the PINN prediction, the DeepONet, the FNO prediction and the corresponding relative error field. c. Comparison of the average performance in terms of nRMSE of the 4DHemoFormer with 3 baseline architectures. Left is for velocity magnitude and right is pressure field. The error bars represent the mean values $\pm$ std., calculated from the method's predictions across the test dataset. *** $p < 0.001$, * $p < 0.05$. d. Comparison of the average performance in terms of correlation coefficients of the 4DHemoFormer with 3 baseline architectures. KDE, Kernel Density Estimation. ICA, internal carotid artery; ECA, external carotid artery; VA, vertebral artery; CCA, Common carotid artery.

Single-step prediction task. In the single-step prediction task, the proposed 4DHemoFormer demonstrated superior performance in reconstructing flow solutions across diverse vessel geometries. Representative results of predicting the velocity field for ICA are presented in Fig. 2b, with corresponding results for the CCA, ECA, and VA provided in Supplementary Figs. 1–3. As presented in Fig. 2c, Fig. 2d, Supplementary Table 2 and Supplementary Fig. 4, 4DHemoFormer achieved superior estimation performance for both pressure field and velocity field (u , v , w , and velocity magnitude) compared with established baselines. Specifically, for velocity magnitude across the complete CFD dataset, 4DHemoFormer yielded a normalized root mean square error (nRMSE) of $1.70 \times 10^{-2} \pm 3.93 \times 10^{-3}$ and a correlation coefficient (CC) of 0.924 ± 0.021 . These metrics represent substantial improvements over baseline models: PINN achieved an nRMSE of $6.80 \times 10^{-2} \pm 2.30 \times 10^{-2}$ with a CC of 0.691 ± 0.179 ; DeepONet

achieved an nRMSE of $6.00 \times 10^{-2} \pm 1.70 \times 10^{-2}$ with a CC of 0.853 ± 0.092 , and FNO achieved an nRMSE of $6.20 \times 10^{-2} \pm 1.50 \times 10^{-2}$ with a CC of 0.839 ± 0.089 (Supplementary Table 2). Furthermore, the proposed model maintained consistent performance advantages across all four vessel sites (Fig. 2c and Supplementary Fig. 4), indicating its high reliability.

Multi-step prediction task. We further evaluated 4DHemoFormer against baseline models for future-state forecasting in the multi-step prediction task, quantifying performance using nRMSE and CC metrics. As illustrated in Fig. 3a, the proposed framework accurately reconstructed the original velocity field given initial conditions, achieving consistently low error across multiple samples, with the ICA serving as the representative example (corresponding results for the CCA, ECA, and VA are provided in Supplementary Figs. 5–7). When averaged across multiple time steps, 4DHemoFormer demonstrated superior estimation performance for velocity components (u , v , w , and velocity magnitude) compared with established baselines over the complete CFD dataset (Fig. 3b, Fig. 3c and Supplementary Fig. 8). Moreover, 4DHemoFormer maintained consistent advantages across all four vessel sites (Fig. 3b), with detailed nRMSE and CC values for all models summarized in Supplementary Table 3. We additionally assessed the temporal error propagation of 4DHemoFormer and baselines given initial conditions. While all models exhibited a gradual increase in estimation error over time, 4DHemoFormer accurately sustained velocity field predictions across extended horizons (Fig. 3d for velocity magnitude and Supplementary Fig. 9 for u , v , w). Compared to established baselines (PINN, DeepONet, and FNO), 4DHemoFormer yielded significantly more accurate and temporally stable pointwise predictions of both velocity magnitude (Fig. 3d) and directional components (Supplementary Fig. 9). Notably, all evaluated models exhibited elevated estimation errors and high variance when predicting the pressure field (Fig. 3e and Supplementary Figs. 8-9), indicative of numerical instability during long-term forecasting. Consequently, extended temporal predictions of pressure dynamics remain broadly challenging across these architectures. Nevertheless, 4DHemoFormer consistently outperformed the baseline models, yielding the most accurate and temporally stable pressure estimates (Fig. 3e).

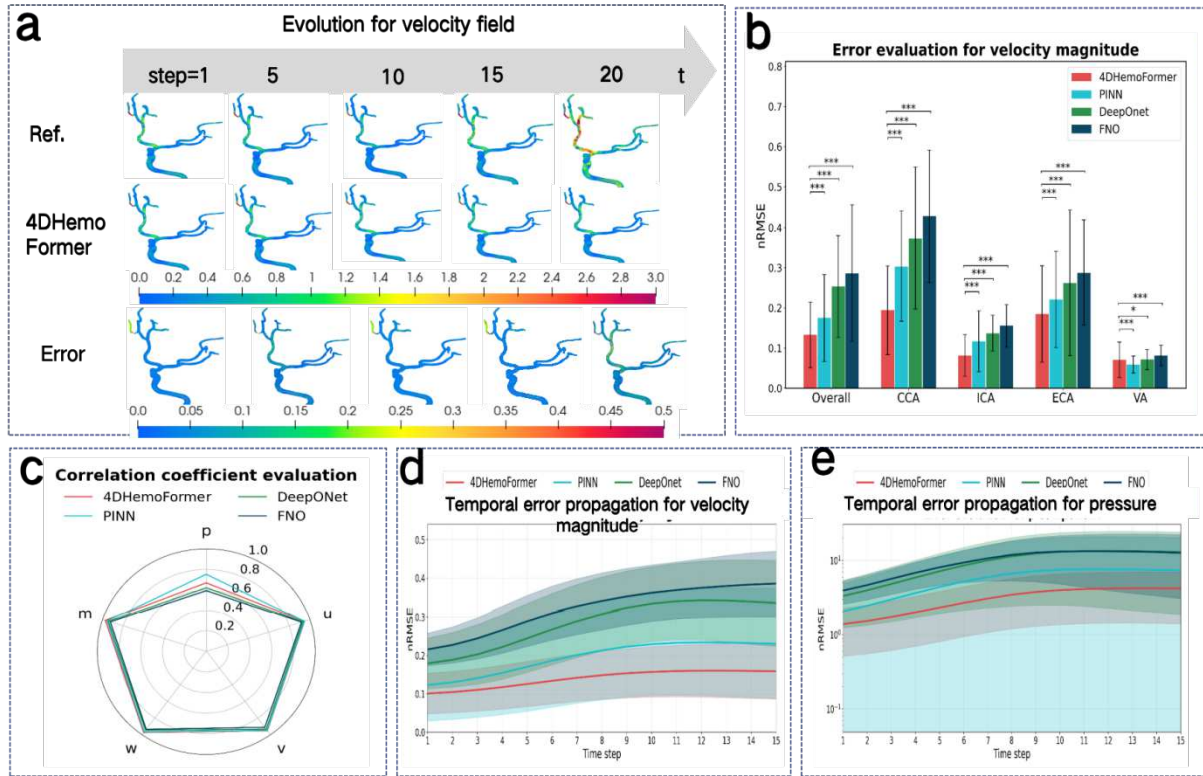


Fig. 3. | Performance of temporal extrapolation in multi-step prediction task. a, Comparison between the hemodynamic velocity fields learned and the reference using ICA as the example. From up to bottom, the panels represent the numerical simulation (denoted as Ref.), the 4DHemoFormer predictions, and the relative error across different time steps. b, Comparison of the average performance for velocity magnitude in terms of nRMSE of the 4DHemoFormer with 3 baseline architectures. The error bars represent the mean values $\pm$ std., calculated from the method's predictions across the test dataset. *** $p < 0.001$. c, Comparison of the average performance in terms of correlation coefficients of the 4DHemoFormer with 3 baseline architectures. d, Temporal error propagation for velocity magnitude predictions in terms of nRMSE. Shaded region represent the variance band with 4DHemoformer and baselines. e, Temporal error propagation for pressure field predictions in terms of nRMSE with 4DHemoformer and baselines. Shaded region represent the variance band. ICA, internal carotid artery; ECA, external carotid artery; VA, vertebral artery; CCA, Common carotid artery.

Cross-geometry generalization on unseen vascular structures

Cross-geometry-shift, arising from variations in vascular structures across subjects or anatomical locations, represent a common challenge in real-world scenarios. However, deep learning models frequently struggle to generalize in out-of-distribution settings where the training data distribution diverges from the target task. Here, we evaluate the generalizability of 4DHemoFormer and baseline methods in cross-subject and cross-vascular contexts. The cross-subject evaluation assesses model generalization to unseen subjects, while the cross-vascular evaluation tests performance on vascular sites distinct from those used during training (Fig. 4a for detail cross-geometry-shift task). As illustrated in Fig. 4b, the data distributions from the training and test datasets differ significantly. We evaluated 4DHemoFormer and

baseline methods under cross-geometry-shift using zero-shot inference. In this protocol, models are applied directly to out-of-distribution test sets without additional supervised learning or fine-tuning, thereby rigorously assessing their intrinsic generalization capabilities.

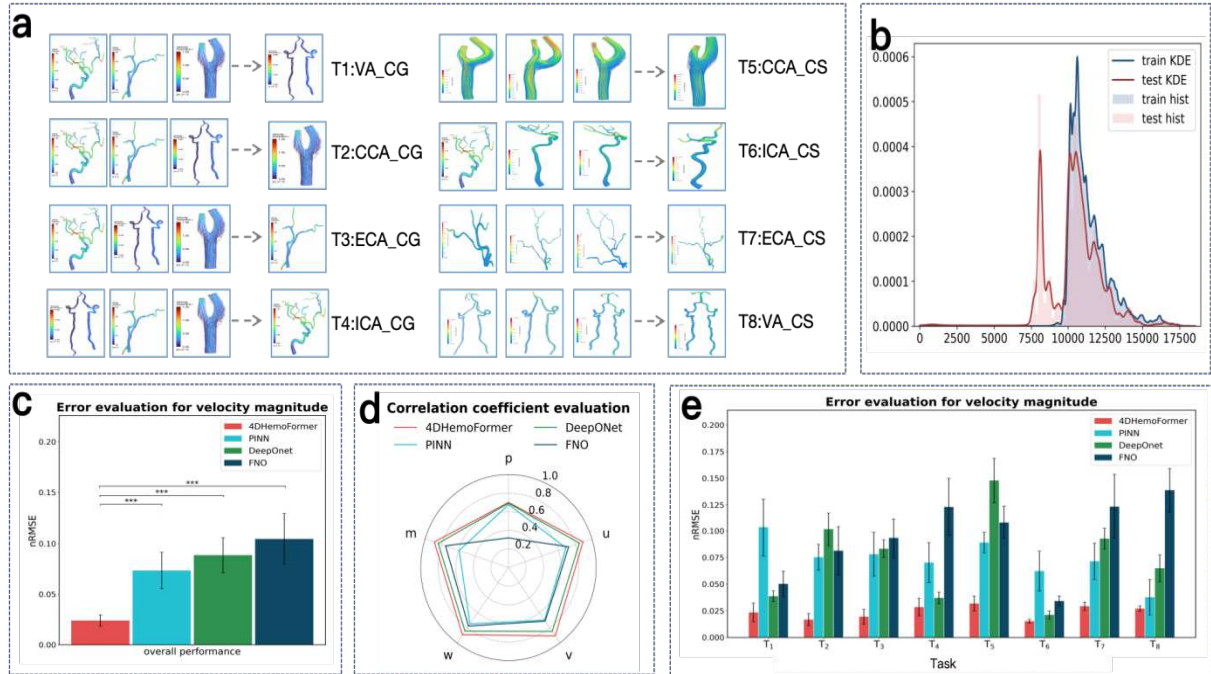


Fig. 4. | Performance for cross-geometry generalization in single-step prediction task. a, Details of tasks. We perform cross-geometry-shift evaluation for cross-vascular (T1-T4) and cross-subject (T5-T8) contexts. The cross-subject evaluation assesses model generalization to unseen subjects, while the cross-vascular evaluation tests performance on vascular sites distinct from those used during training. b, Data distribution in the training and testing dataset, with the pressure component as the example. c, Overall performance evaluation for velocity magnitude of 4DHemoFormer and baselines on 8 cross-geometry tasks in terms of nRMSE. The bar graph indicates the mean $\pm$ std. *** $p < 0.001$. d, Overall performance evaluation of 4DHemoFormer and other models on 8 cross-geometry tasks in terms of correlation coefficient. e, Detailed evaluation of 4DHemoFormer and baselines for velocity magnitude on each cross-geometry task. KDE, Kernel Density Estimation.

Single-step prediction task. As quantified in Fig. 4c, Fig. 4d, Supplementary Fig. 10 and Supplementary Table 4, 4DHemoFormer demonstrated superior estimation performance for both pressure and velocity components (u , v , w , and velocity magnitude) compared to established baselines. Specifically, across all of the cross-geometry-shift tasks for velocity magnitude prediction, 4DHemoFormer achieved an average nRMSE of $2.39 \times 10^{-2} \pm 0.56 \times 10^{-2}$ and CC of 0.890 ± 0.051 . These metrics represent substantial improvements over baseline models: PINN (nRMSE: $7.34 \times 10^{-2} \pm 1.81 \times 10^{-2}$; CC: 0.5491 ± 0.149), DeepONet (nRMSE: $8.85 \times 10^{-2} \pm 1.74 \times 10^{-2}$; CC: 0.843 ± 0.144), and FNO (nRMSE: $10.46 \times 10^{-2} \pm 2.46 \times 10^{-2}$; CC: 0.755 ± 0.096) (Supplementary Table 4). Furthermore, the proposed model maintained

consistent accuracy across all cross-geometry-shift tasks, highlighting the robust generalizability of 4DHemoFormer to unseen vascular geometries and diverse subjects (Fig. 4e for velocity magnitude; Supplementary Fig. 11 for pressure and individual velocity components).

Multi-step prediction task. We further evaluated our method against baseline models for future-state forecasting in the multi-step prediction task. When averaged across multiple time steps, 4DHemoFormer demonstrated superior estimation performance in terms of nRMSE for velocity field (u , v , w , and velocity magnitude) compared with established baselines across all of the cross-geometry-shift tasks (Fig. 5a for velocity magnitude; Supplementary Fig. 12 for pressure and individual velocity components). Moreover, 4DHemoFormer maintained consistent advantages across most cross-geometry-shift task (Fig. 5c for velocity magnitude; Supplementary Fig. 13 for pressure and individual velocity components), with detailed nRMSE and CC values for all models summarized in Supplementary Table 5. By further analyzing the temporal evolution of prediction errors (Fig. 5d, Fig. 5e and Supplementary Fig. 14), we can clearly observe the inherent error propagation dynamics typical of multi-step forecasting. For the velocity field (Fig. 5d for velocity magnitude; Supplementary Fig. 14 for individual velocity components), the error across all models exhibits a gradually upward. Despite the inevitable accumulation of errors, 4DHemoFormer demonstrates superior error mitigation capabilities. It not only achieves the lowest initial error but also displays the most gradual growth rate, ultimately stabilizing at a level significantly lower than all established baselines (PINN, DeepONet, and FNO). Furthermore, the narrow and uniform variance band (shaded region) associated with 4DHemoFormer indicates high precision and robust consistency across different samples during extended velocity forecasting. Conversely, error propagation in the pressure field (Fig. 5e) reveals a much more severe and complex pattern. During long-term forecasting, the mean errors of the baseline models (particularly FNO and DeepONet) escalate rapidly. While the PINN model achieves lower errors for certain samples, its variance band exhibits extreme divergence. This massive variance strongly indicates that the baseline models suffer from severe numerical instability when modeling the long-term temporal evolution of pressure dynamics. Although long-term pressure forecasting remains inherently challenging across all evaluated architectures, 4DHemoFormer effectively mitigates this catastrophic error

accumulation. Overall, 4DHemoFormer consistently maintains the lowest mean error, and crucially, its variance is strictly confined to a narrow range, entirely avoiding the drastic divergence observed in the baselines. These findings demonstrate that 4DHemoFormer successfully suppresses the cascading amplification of errors, providing the most reliable and temporally stable long-term predictions, even when forecasting highly volatile pressure fields.

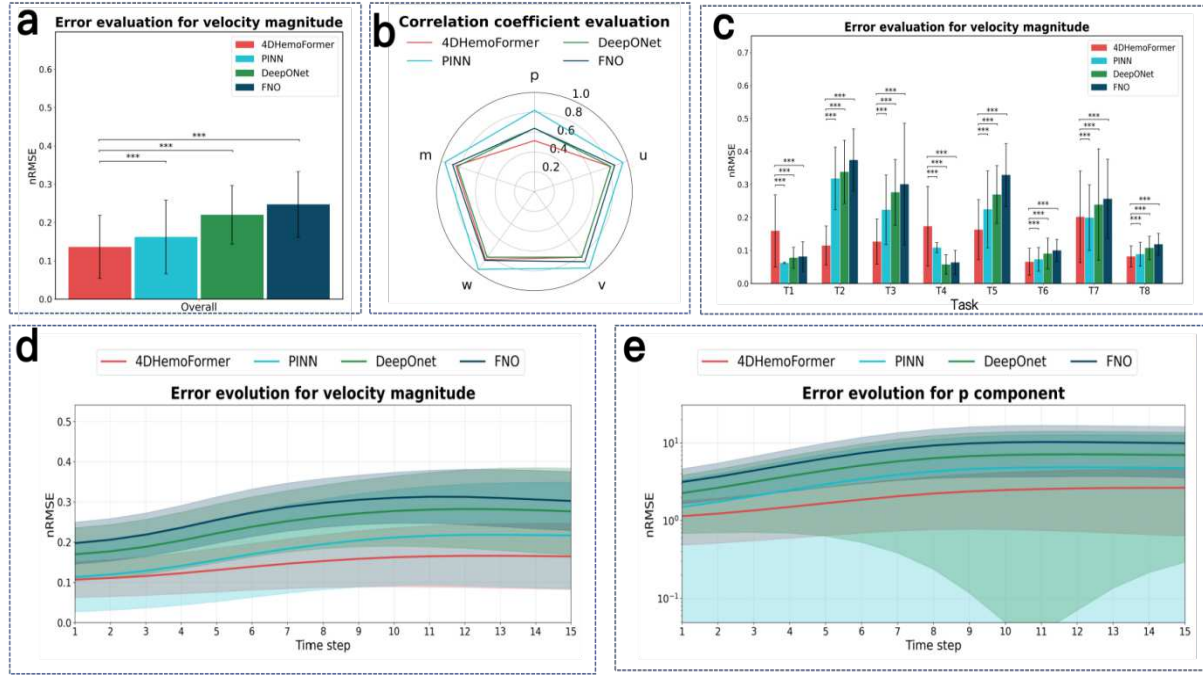


Fig. 5. | Performance for cross-geometry generalization in multi-step prediction task. a, Overall performance evaluation of 4DHemoFormer and baselines for velocity magnitude on 8 cross-geometry tasks in terms of nRMSE. The bar graph indicates the mean $\pm$ std. *** $p < 0.001$. b, Overall performance evaluation of 4DHemoFormer and baselines on 8 cross-geometry tasks in terms of correlation coefficient. c, Detailed evaluation of 4DHemoFormer and baselines for velocity magnitude on each cross-geometry task. d, Temporal error propagation for velocity magnitude predictions with 4DHemoformer and baselines. Shaded region represent the variance band. e, Temporal error propagation for pressure field predictions with 4DHemoformer and baselines. Shaded region represent the variance band.

Cross-domain transferability from *in silico* simulations to *in vivo* data

To further evaluate the clinical applicability of 4DHemoFormer in real-world settings, we investigated its capacity to learn spatio-temporal dynamics using our In-vivo dataset. Zero-shot and data-efficient performance evaluation are both considered (see performance evaluation in Supplementary Table 6).

Zero-shot inference. For zero-shot inference, we train and validate the model on the CFD dataset and test its performance directly on the *in vivo* dataset. Fig. 6a presents the visualization

of the predicted flow fields versus the ground truth in the multi-step prediction task, indicating that 4DHemoFormer captures the fundamental spatio-temporal hemodynamics well even without seeing real-world data. We systematically compared performance across temporal extrapolation, cross-geometry generalization, and cross-domain transferability. As expected, a progressive degradation in performance is observed as the generalization task becomes more complex (Fig. 6 b-e). The model performs best in temporal generalization, consistently achieving the lowest nRMSE. When subjected to cross-geometry generalization, the prediction error increases slightly, indicating that while the model retains strong predictive capacity, shifting to a new spatial domain introduces measurable discrepancies. Naturally, cross-domain transferability (sim-to-real) presents the most rigorous challenge, exhibiting the highest nRMSE and weakest correlation across metrics. This baseline performance transparently reflects the profound intrinsic difficulties of applying models trained on idealized synthetic data to complex, real-world in vivo measurements.

Data-efficient generalizability. To demonstrate how our framework practically overcomes the aforementioned "sim-to-real" gap, we investigated the data efficiency of 4DHemoFormer by fine-tuning it on merely 10% of the subjects in the in vivo dataset. This setup mimics real-world clinical scenarios where labeled MRI data is notoriously scarce. Notably, with this highly limited fine-tuning data, 4DHemoFormer exhibits a drastic performance improvement over the zero-shot baseline. It achieves an nRMSE of $3.51 \times 10^{-2} \pm 8.82 \times 10^{-3}$ for the single-step prediction task (down from zero-shot: $6.28 \times 10^{-2} \pm 9.12 \times 10^{-3}$) and $7.76 \times 10^{-2} \pm 1.25 \times 10^{-2}$ for the multi-step task (down from zero-shot: $1.04 \times 10^{-1} \pm 2.95 \times 10^{-2}$) (Fig. 6b and 6d). Crucially, this rapid adaptation validates our core hypothesis: 4DHemoFormer successfully bridges the "sim-to-real" gap. By effectively transferring the rich physical priors learned from the CFD domain, the model requires only minimal real-world data to calibrate its predictions, thereby providing a highly data-efficient and clinically viable solution for personalized hemodynamic modeling.

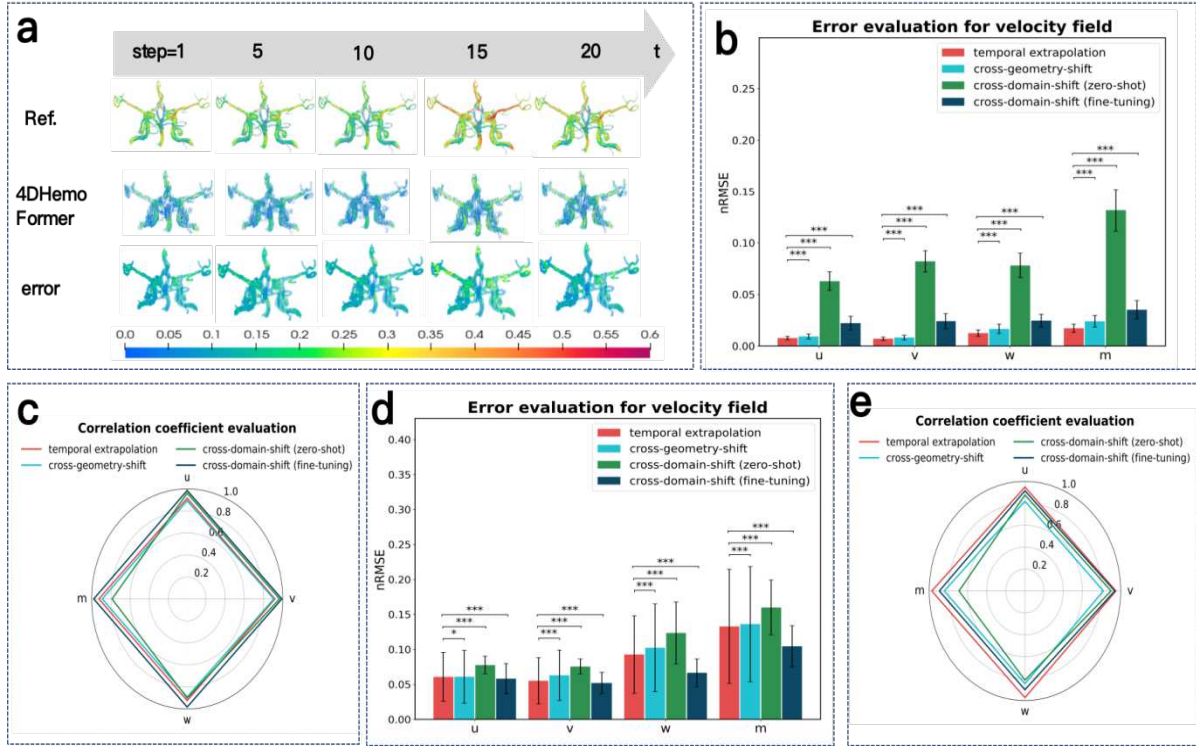


Fig. 6. | Performance of 4DHemoFormer on cross-domain transferability to *in vivo* data. a, Visualization of velocity field between the hemodynamic fields learned with zero-shot inference and the reference in multi-step prediction task. From up to bottom, the panels represent the 4D-flow MRI collection (denoted as Ref.), the 4DHemoFormer prediction, and the relative error with time steps. b, Comparison of the performance on temporal extrapolation, cross-geometry generalization and cross-domain transferability (zero-shot and fine-tuning on 10% data) in single-step prediction task in terms of nRMSE. The error bars represent the mean values $\pm$ std., calculated from the method's predictions across the test dataset. *** $p < 0.001$. c, Comparison of the performance on temporal extrapolation, cross-geometry generalization and cross-domain transferability (zero-shot and fine-tuning on 10% data) in single-step prediction task in terms of correlation coefficient. d, Comparison of the performance on temporal extrapolation, cross-geometry generalization and cross-domain transferability (zero-shot and fine-tuning on 10% data) in multi-step prediction task in terms of nRMSE. The error bars represent the mean values $\pm$ std., calculated from the method's predictions across the test dataset. *** $p < 0.001$. * $p < 0.05$. e, Comparison of the performance on temporal extrapolation, cross-geometry generalization and cross-domain transferability (zero-shot and fine-tuning on 10% data) in multi-step prediction task in terms of correlation coefficient.

Performance evaluation in pathological conditions (Stenosis)

To evaluate the model's robustness to pathological geometric alterations, we compared its predictive accuracy between normal and stenosed vascular cohorts on *in vivo* data (Supplementary Fig. 15). The model maintained high predictive fidelity across both states, demonstrating a robust capacity to resolve the complex, disturbed flow patterns induced by arterial narrowing. Quantitative analysis revealed no performance degradation in the presence of pathology; rather, the stenosed group yielded accuracy metrics highly consistent with those of the healthy cohort. For instance, in multi-step prediction tasks, the model achieved comparable performance for velocity magnitude across both cohorts ($\text{nRMSE: } 1.32 \times 10^{-1} \pm$

2.06 $\times 10^{-2}$ for the normal group vs. 1.27 $\times 10^{-1} \pm 1.24 \times 10^{-2}$ for the stenosed group). Similar robust performances are observed across the individual velocity components, demonstrating the model's reliability in simulating disturbed hemodynamics without experiencing significant error accumulation.

DISCUSSION

Despite recent progress in physics-informed AI for fluid dynamics²¹, the successful clinical translation of these models for real-world cerebrovascular hemodynamics assessment remains severely hindered by the formidable "sim-to-real" gap. To bridge this divide, our work introduces a generalizable, physics-informed AI framework. We first establish 4DHemoDB, a pioneering cross-domain database pairing high-fidelity *in silico* simulations with *in vivo* 4D-flow MRI measurements. This dataset not only validates our framework but serves as a foundational benchmark for downstream clinical transition tasks, such as hemodynamic super-resolution. To enhance the adaptability of AI-based models in clinical environments, we developed 4DHemoFormer, a physics-informed generalist Transformer. Through a rigorous evaluation encompassing temporal extrapolation, cross-geometry generalization and cross-domain transferability, 4DHemoFormer demonstrates unprecedented robustness when deployed on unseen vascular structures and clinical data. This capability enables the rapid, high-fidelity reconstruction of patient-specific hemodynamics, paving the way for proactive interventions in neurovascular diseases. However, realizing the full clinical potential of this framework requires a nuanced understanding of its underlying mechanisms, inherent trade-offs, and current boundaries.

The superior performance and robust stability of 4DHemoFormer in predicting both velocity and pressure fields can be attributed to its architectural design. Unlike traditional purely data-driven architectures, the physics-biased self-attention mechanism within our Transformer explicitly embeds fluid dynamic constraints into the information routing process. This physical inductive bias actively suppresses non-physical spurious correlations, enabling the model to capture the complex, long-range spatio-temporal dependencies inherent in pulsatile blood flow while strictly maintaining physical consistency. Furthermore, the integration of physics-informed constraints acts as a regularizer, ensuring that the network's predictions strictly adhere

to the Navier-Stokes equations. This dual approach significantly mitigates the numerical instability and physical inconsistency often observed in purely data-driven models. A crucial challenge introduced by the embedding of physical constraints is computational overhead problem. Incorporating the Navier-Stokes equations into the loss function requires computing high-order spatio-temporal derivatives of the network's outputs with respect to its inputs via automatic differentiation. In a 4D spatio-temporal domain, this process significantly amplifies the computational overhead during the training phase, demanding substantial GPU memory and extending convergence times compared to purely data-driven baselines (a list of computational costs for each model is shown in Supplementary Table 7, with the corresponding runtime environment stated in Supplementary Note 6). However, this training overhead does not compromise deployment efficiency. During inference, 4DHemoFormer demonstrates remarkable computational speed, effectively overcoming the prohibitive time barriers associated with traditional CFD simulations and in vivo data acquisition. For instance, while standard 4D flow MRI acquisition takes 575 seconds and traditional numerical simulations require 415.43 seconds of computation per simulated second, 4DHemoFormer completes inference in just 23.5 seconds (Supplementary Table 7). This massive reduction in turnaround time bypasses traditional computational bottlenecks, paving the way for real-time surgical planning and rapid, point-of-care clinical decision-making. Furthermore, the performance of the model is sensitive to the selection of the physics loss weight. Taking the temporal extrapolation task as an example, our results indicate that a weight of 0.5 optimally balances physical constraints and prediction accuracy (Supplementary Fig. 16). Therefore, while the physics-informed approach is indispensable for ensuring the robustness and generalizability of hemodynamic forecasting, future iterations must explore adaptive loss-weighting strategies and more memory-efficient derivative computation methods to mitigate these computational bottlenecks without sacrificing physical fidelity.

In evaluating the multi-step prediction task, we observed an apparent inconsistency between distance-based metrics (nRMSE) and correlation-based metrics (CC). While our model exhibits a marginally lower CC compared to the baseline, it achieves a notably smaller nRMSE. In the context of clinical hemodynamic modeling, we emphasize that nRMSE is the more critical evaluation metric. CC primarily measures linear correlation and is insensitive to

absolute magnitude shifts; a model could theoretically yield a high CC while producing physically unrealistic values due to scaling errors. In contrast, nRMSE directly quantifies absolute point-wise accuracy. Since precise absolute quantification of flow metrics (e.g., velocity, pressure) is essential for reliable clinical assessment, the substantial reduction in nRMSE demonstrates the superior physical fidelity of our model.

Despite this superior point-wise accuracy, a closer examination of transient hemodynamic events reveals a common challenge in deep learning-based fluid solvers. While 4DHemoFormer successfully tracks the global macroscopic trends and the phase of the pulsatile flow, it exhibits a tendency to underestimate peak magnitudes during highly dynamic phases (e.g., peak systole) or in regions with steep spatial gradients. From a clinical perspective, this smoothing effect is non-trivial, as peak velocities directly dictate critical biomechanical markers, such as wall shear stress, which are essential for evaluating aneurysm rupture risk or plaque vulnerability. Algorithmically, this observed smoothing effect can be attributed to two primary algorithmic factors. First, neural networks inherently suffer from spectral bias, prioritizing the learning of low-frequency macroscopic flow patterns over high-frequency transient details^{54,55}. Second, the reliance on standard regression losses (e.g., Mean Squared Error) often induces a statistical smoothing effect⁵⁶. Consequently, to minimize the overall global error and maintain stability during long-term multi-step extrapolation, the model implicitly introduces numerical dissipation, which dampens extreme values at sharp spatio-temporal gradients.

Beyond the smoothing of high-frequency velocity features, a critical finding of our multi-step evaluation is the ubiquitous difficulty in accurately forecasting the long-term temporal evolution of the pressure field. Compared to the velocity field, pressure dynamics are highly sensitive to boundary conditions and are implicitly coupled with velocity through the continuity equation²³. In a multi-step rollout, even marginal errors in early velocity or pressure predictions are fed back into the model, leading to catastrophic error accumulation over time. Furthermore, the lack of absolute pressure references in localized vessel segments exacerbates this numerical drift, making long-term pressure stability a formidable challenge. To fully close the loop on long-term hemodynamic forecasting, future work will focus on incorporating more rigorous physiological boundary condition formulations, exploring higher-order physical constraints to

preserve high-frequency kinematics, and designing advanced temporal rollout strategies to effectively arrest error accumulation.

Looking forward, the clinical implications of our framework are manifold. By bridging the sim-to-real gap, 4DHemoFormer paves the way for non-invasive, patient-specific hemodynamic profiling using routine clinical imaging. In neurovascular disease management, this capability could revolutionize risk stratification, such as predicting the rupture risk of intracranial aneurysms or assessing stroke vulnerability, by providing high-fidelity flow dynamics without the need for invasive catheterization. Furthermore, its robust cross-geometry generalization holds significant promise for pre-operative treatment planning. Clinicians could leverage this model to simulate post-operative hemodynamics, optimizing the selection and placement of flow diverters or stents tailored to individual patient anatomies. Ultimately, integrating such physics-informed AI into routine diagnostic workflows could accelerate the transition toward precision cerebrovascular medicine.

Methods

The 4DHemoDB

Establishment of 4DHemoDB. To construct the 4DHemoDB, we enrolled 40 subjects (subject demographics are summarized in Supplementary Table 1). The data collection protocol (detailed in Supplementary Note 1), involved three primary acquisitions for each subject: (1) a 3D Time-of-Flight (TOF) MRI scan to reconstruct vascular geometry; (2) Doppler ultrasonography to measure bilateral common and internal carotid artery velocity waveforms, which served as patient-specific inflow boundary conditions; and (3) a 4D Flow MRI scan to capture *in vivo* hemodynamic data. All MRI sequences were acquired using a 3.0 T scanner (uMR Omega, United Imaging Healthcare, Shanghai, China) at the United Imaging iHealthcare Center.

The generation of patient-specific CFD data utilized a multi-stage workflow to transform clinical images into high-fidelity, time-resolved flow fields. Key steps included geometry reconstruction, mesh generation, assignment of fluid properties and boundary conditions, numerical simulation, and data extraction (see Supplementary Note 2 and Supplementary Fig.

17). For *in vivo* data analysis, the acquired images were processed using commercial 4D Flow data post-processing software (iTFlow, Cardio Flow Design Inc., Japan) to quantify spatio-temporal hemodynamics (see Supplementary Note 3 and Supplementary Fig. 18).

This study was conducted at the Medical Imaging Diagnostic Center, Shanghai IMAGINGRACE and was approved by the Institutional Review Board of Fuwai Hospital Shenzhen, Chinese Academy of Medical Sciences (Approval number: SP2024095(01)). Informed written consent was obtained from all participants.

Datasets. To develop and validate the models, the CFD dataset of 4DHemoDB was reorganized into 3 sets (Extended Data Fig. 2a), including one in-distribution (ID) set and two out-of-distribution (OOD) sets. The ID set, comprising four subsets corresponding to the CCA, ICA, ECA, and VA, was used to evaluate temporal extrapolation performance. Within each subset, the data were partitioned along the temporal dimension. Specifically, for each case, the initial 9 seconds (180 time steps) were allocated for model development and further divided into training and validation sets at a 7:2 ratio. The final 1 second (the last 20 time steps) of each case was strictly held out to test the model's temporal extrapolation capabilities. The cross-vascular OOD set was used to assess generalization across different vessel types (corresponding to cross-geometry-shift task T1-T4). For this set, a leave-one-vessel-out cross-validation strategy was employed. In each iteration, one of the four vascular subsets (e.g., CCA) was reserved entirely as the unseen test set for geometric generalization, while the remaining three subsets (e.g., ICA, ECA, VA) were utilized for training and validation. The cross-subject OOD set was used to evaluate patient-specific generalization (corresponding to cross-geometry-shift task T5-T8), the cohort of 40 subjects was randomly partitioned at the subject level into mutually exclusive training, validation, and test sets at a ratio of 7:2:1. This ensures that the morphological variations of the test subjects were completely unseen during the training phase. To further evaluate the clinical applicability of our model, we leveraged the In-Vivo subset of 4DHemoDB to validate its cross-domain transfer capabilities. Specifically, the entire CFD dataset was utilized for model development and internal validation, whereas the *in vivo* data served as an independent test set for evaluating performance under domain shift.

The 4DHemoFormer

The core of 4DHemoFormer consists of a deep stack of Physics-Biased Transformer blocks and a dual-way loss formulation, which are specifically designed to embed fluid dynamics constraints directly into the network's forward pass and optimization landscape, respectively.

Physics-Bias Attention Mechanism. A fundamental limitation of standard self-attention in modeling physical systems is its purely data-driven, fully connected nature, which ignores the local topological and physical connectivity of the spatial domain. To address this, we introduce a Physics-Biased Attention Mechanism. We explicitly inject a Physics-Bias matrix into the attention logits prior to the Softmax activation. This bias matrix acts as a strong architectural inductive bias, derived from the physical priors of the governing fluid equations. Specifically, it encodes the spatial connectivity defined by the Diffusion Bias (approximating the Laplacian operator) and the Convection Bias (approximating the advection operator). Mathematically, the modified attention operation is formulated as:

$$\text{Attention}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \text{Softmax}\left(\frac{\mathbf{Q}\mathbf{K}^T}{\sqrt{d_k}} + \mathbf{M}_{\text{physics}}\right)\mathbf{V} \quad (6)$$

where $\mathbf{M}_{\text{physics}}$ is the Physics-Bias matrix combined with diffusion bias and convection bias.

To encapsulate the isotropic nature of viscous momentum dissipation governed by the Laplacian operator, we formulate the diffusion bias matrix $\mathbf{M}_{\text{diffusion}}$. For any given spatial node i and its neighbor j , the bias is computed using a Gaussian radial basis function based on their Euclidean distance, which can be expressed as

$$\mathbf{M}_{\text{diffusion}}[i, j] = \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|_2^2}{2\sigma^2}\right) \quad (7)$$

where $\mathbf{x}_i$ and $\mathbf{x}_j$ are the 3D spatial coordinates for i and j , respectively, σ is a learnable or predefined length-scale parameter controlling the diffusion radius. This design enforces a distance-aware connectivity that strictly mirrors the physical behavior of fluid internal friction.

The advection of momentum, represented by the non-linear term $(\mathbf{u} \cdot \nabla)\mathbf{u}$, dictates an anisotropic, flow-directed propagation of information. To embed this physical causality into the attention mechanism, we introduce the convection bias matrix $\mathbf{M}_{\text{convection}}$ inspired by the upwind scheme in computational fluid dynamics. In detail, we mathematically define the unstream-downstream relationship between spatial nodes. Let i be the query node and j be the key node, the convection bias matrix can be expressed as

$$M_{convection}[i,j] = \begin{cases} 0, & \text{if } \mathbf{u}_j \cdot \frac{\mathbf{x}_i - \mathbf{x}_j}{\|\mathbf{x}_i - \mathbf{x}_j\|_2 + \epsilon} \geq 0 \\ -\infty, & \text{if } \mathbf{u}_j \cdot \frac{\mathbf{x}_i - \mathbf{x}_j}{\|\mathbf{x}_i - \mathbf{x}_j\|_2 + \epsilon} < 0 \end{cases} \quad (8)$$

where $\mathbf{x}_i$ and $\mathbf{x}_j$ are the 3D spatial coordinates for i and j , $\mathbf{u}_j$ is the velocity vector for j . This formula ensures that attention is exclusively amplified when fluid flows from j towards i , while downstream-to-upstream information leakage is suppressed. By adding this matrix, the information routing between tokens is explicitly modulated by the underlying fluid transport phenomena, ensuring that the network intrinsically respects local physical interactions rather than relying solely on spurious data correlations.

Physics-Informed Loss Landscape and Optimization. To further ensure the predicted velocity field and pressure field adhere to the fundamental laws of fluid mechanics, the network is optimized using a composite loss function that simultaneously minimizes data discrepancies and physical equation residuals. We leverage Automatic Differentiation to compute the exact spatial and temporal derivatives of the network's outputs with respect to its inputs. This allows us to formulate the Physics Loss ($\mathcal{L}_{physics}$) based on the incompressible Navier-Stokes (NS) equations, which comprises mass conservation and momentum conservation components. Mass conservation penalizes the divergence of the velocity field to ensure fluid incompressibility, which can be expressed as:

$$\mathcal{L}_{mass} = \|\nabla \cdot \hat{\mathbf{U}}\|_2 \quad (9)$$

Where $\hat{\mathbf{U}}$ is the predicted velocity field for x -direction velocity $u \in \mathbb{R}$, y -direction velocity $v \in \mathbb{R}$, and z -direction velocity $w \in \mathbb{R}$. Momentum conservation penalizes the residual of the momentum equations, accounting for temporal acceleration, convective acceleration, pressure gradients, and viscous diffusion, expressing as:

$$\mathcal{L}_{mom} = \left\| \frac{\partial \hat{\mathbf{U}}}{\partial t} + (\hat{\mathbf{U}} \cdot \nabla) \hat{\mathbf{U}} + \frac{1}{\rho} \nabla \hat{p} - \nu \nabla^2 \hat{\mathbf{U}} \right\|_2 \quad (10)$$

Concurrently, a Data Loss ($\mathcal{L}_{data}$) is computed to anchor the predictions to the available ground truth measurements. The overall optimization objective is defined by the Total Loss ($\mathcal{L}_{total}$), which balances the physical constraints and data fidelity via a weighting hyperparameter α , which can be expressed as

$$\mathcal{L}_{total} = (1 - \alpha) \mathcal{L}_{data} + \alpha (\mathcal{L}_{mass} + \mathcal{L}_{mom}) \quad (11)$$

By navigating this explicitly constrained loss landscape, the model avoids non-physical local minima, thereby achieving high-fidelity hemodynamic predictions with enhanced generalizability, even in data-sparse regimes.

Implemental details and Parameters are presented in Supplementary Note 4 and Supplementary Table 8.

Evaluation metrics

nRMSE⁵⁷ is selected as the primary evaluation metric in our study, defined as

$$nRMSE(\mathbf{x}, \hat{\mathbf{x}}) = \frac{\|\mathbf{x} - \hat{\mathbf{x}}\|_2}{\mathbf{x}_{\max} - \mathbf{x}_{\min}}$$

$$(12)$$

where $\mathbf{x}$ is the ground-truth data and $\hat{\mathbf{x}}$ is the predicted data. In this context, a lower nRMSE indicates a higher pointwise accuracy. We also use the Pearson correlation coefficient to show the correlation between the reference and the predictions, defined as

$$Correlation(\mathbf{x}, \hat{\mathbf{x}}) = \frac{(\mathbf{x} - \mathbf{x}_{mean}) \cdot (\hat{\mathbf{x}} - \hat{\mathbf{x}}_{mean})}{\|\mathbf{x} - \mathbf{x}_{mean}\|_2 \|\hat{\mathbf{x}} - \hat{\mathbf{x}}_{mean}\|_2} \quad (13)$$

where $\mathbf{x}$, $\hat{\mathbf{x}}$ are two vectors with the same dimension, and the subscript ‘mean’ refers to the mean value of all the entries of the vector. The Pearson correlation coefficient is a value between -1 and 1, and a higher value indicates a stronger correlation between these two vectors.

For the relative error field reported in Fig. 2b, Fig. 3a, Fig. 6a, each entry of the error field is defined as

$$r_i = |x_i - \hat{x}_i| \quad (14)$$

where x_i and $\hat{x}_i$ represent the i th entries of the ground-truth data $\mathbf{x}$ and predicted data $\hat{\mathbf{x}}$, respectively, and $|\cdot|$ denotes the absolute value of the input.

Baseline Methods

We consider three different baseline models, namely, PINN^{21,22}, DeepONets³³ and FNO³⁴. PINN is a famous deep learning paradigm that integrates the residual of physical laws into the loss function as a regularizer or soft constraint. By minimizing such a regularized loss, it can handle many forward and inverse analyses of partial differential equation systems. DeepONets

leverage the universal operator approximation theorem to design a network architecture that can learn the mapping between function spaces³³, while they may not perform well in systems with very complicated spatiotemporal patterns due to the limited model capacity. FNO attempts to map between function spaces, leveraging the global information of Fourier features³⁴. Hemodynamic data, including CFD data and In-vivo data are derived on complex, patient-specific vascular geometries, is typically defined on unstructured grids. The spatial points are irregularly distributed to efficiently conform to intricate anatomical shapes. However, algorithms like FNO are fundamentally designed for data on structured, uniform grids. To bridge the domain gap between the irregular, non-uniform nature of real-world biomedical data and the structural requirements of powerful mesh-based neural operator algorithms, a voxelization strategy is deployed on the dataset before modeling with FNO and DeepONet (Pipeline presented in the Supplementary Note 5). These baseline methods are compared for both single-step and multi-step prediction tasks, trained in an end-to-end manner that adopts the observation data as input. Parameter settings of these baseline methods are listed in Supplementary Table 8.

Experimental Environments

All training and evaluation procedures were conducted on a dedicated Linux server (Ubuntu 24.04.3 LTS) equipped with an Intel Xeon Platinum 8468 CPU and a single NVIDIA H100 GPU (Driver Version: 550.163.01, CUDA Version: 12.4).

References:

- [1] Cooper, L. L. et al. Components of hemodynamic load and cardiovascular events: the Framingham Heart Study. *Circulation* 131, 354-361 (2015).
- [2] Nixon, A. M., Gunel, M. & Sumpio, B. E. The critical role of hemodynamics in the development of cerebral vascular disease: a review. *J. Neurosurg.* 112, 1240-1253 (2010).
- [3] Tsivgoulis, G. & Alexandrov, A. V. Cerebral hemodynamics in acute stroke: pathophysiology and clinical implications. *J. Vasc. Interv. Neurol.* 1, 65 (2008).
- [4] Cebal, J. R., Castro, M. A., Burgess, J. E., Pergolizzi, R. S., Sheridan, M. J. & Putman, C. M. Characterization of cerebral aneurysms for assessing risk of rupture by using patient-specific computational hemodynamics models. *Am. J. Neuroradiol.* 26, 2550-2559 (2005).
- [5] Cebal, J. R., Mut, F., Raschi, M., Scrivano, E., Ceratto, R., Lylyk, P. & Putman, C. M. Aneurysm rupture following treatment with flow-diverting stents: computational hemodynamics analysis of treatment. *Am. J. Neuroradiol.* 32, 27-33 (2011).
- [6] Pour-Ghaz, I. et al. Accuracy of non-invasive and minimally invasive hemodynamic monitoring: where do we stand?. *Ann. Transl. Med.* 7, 421 (2019).
- [7] Markl, M., Frydrychowicz, A., Kozerke, S., Hope, M. & Wieben, O. 4D flow MRI. *J. Magn. Reson. Imaging* 36, 1015-1036 (2012).

- [8] Garcia, J., Barker, A. J. & Markl, M. The role of imaging of flow patterns by 4D flow MRI in aortic stenosis. *JACC Cardiovasc. Imaging* 12, 252-266 (2019).
- [9] Patel, N. M. et al. Super-resolving and denoising 4D flow MRI of neurofluids using physics-guided neural networks. *Ann. Biomed. Eng.* 53, 331-347 (2025).
- [10] Pontone, G. & Rabbat, M. G. The new era of computational fluid dynamics in CT angiography: far beyond the FFR number. *JACC Cardiovasc. Imaging* 10, 674-676 (2017).
- [11] Mittal, R. et al. Computational modeling of cardiac hemodynamics: Current status and future outlook. *J. Comput. Phys.* 305, 1065-1082 (2016).
- [12] Morris, P. D. et al. Computational fluid dynamics modelling in cardiovascular medicine. *Heart* 102, 18-28 (2016).
- [13] Xiang, J., Tutino, V. M., Snyder, K. V. & Meng, H. CFD: computational fluid dynamics or confounding factor dissemination? The role of hemodynamics in intracranial aneurysm rupture risk assessment. *AJNR Am. J. Neuroradiol.* 35, 1849-1857 (2014).
- [14] Zienkiewicz, O. C. & Taylor, R. L. The finite element method set (Elsevier, 2005).
- [15] Eymard, R., Gallouët, T. & Herbin, R. Finite volume methods. *Handb. Numer. Anal.* 7, 713-1018 (2000).
- [16] Zawawi, M. H. et al. A review: Fundamentals of computational fluid dynamics (CFD). *In AIP Conf. Proc.* 2030, 020252 (AIP Publishing, 2018).
- [17] Zhang, J. M. et al. Perspective on CFD studies of coronary artery disease lesions and hemodynamics: a review. *Int. J. Numer. Method. Biomed. Eng.* 30, 659-680 (2014).
- [18] Brunton, S. L. & Kutz, J. N. Promising directions of machine learning for partial differential equations. *Nat. Comput. Sci.* 4, 483-494 (2024).
- [19] Kochkov, D. et al. Machine learning - accelerated computational fluid dynamics. *Proc. Natl. Acad. Sci. U.S.A.* 118, e2101784118 (2021).
- [20] Cheung, K. C. & See, S. Recent advance in machine learning for partial differential equation. *CCF Trans. High Perform. Comput.* 3, 298-310 (2021).
- [21] Karniadakis, G. E. et al. Physics-informed machine learning. *Nat. Rev. Phys.* 3, 422-440 (2021).
- [22] Raissi, M., Perdikaris, P. & Karniadakis, G. E. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *J. Comput. Phys.* 378, 686-707 (2019).
- [23] Raissi, M. Deep hidden physics models: Deep learning of nonlinear partial differential equations. *J. Mach. Learn. Res.* 19, 1-24 (2018).
- [24] Raissi, M., Yazdani, A. & Karniadakis, G. E. Hidden fluid mechanics: Learning velocity and pressure fields from flow visualizations. *Science* 367, 1026 - 1030 (2020).
- [25] Belbute-Peres, F. D. A., Economou, T. & Kolter, Z. Combining differentiable PDE solvers and graph neural networks for fluid flow prediction. *In Int. Conf. Mach. Learn.* (PMLR, 2020).
- [26] Raissi, M., Perdikaris, P. & Karniadakis, G. E. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *J. Comput. Phys.* 378, 686-707 (2019).
- [27] Jagtap, A. D., Kharazmi, E. & Karniadakis, G. E. Conservative physics-informed neural networks on discrete domains for conservation laws: Applications to forward and inverse problems. *Comput. Methods Appl. Mech. Eng.* 365, 113028 (2020).
- [28] Jagtap, A. D. & Karniadakis, G. E. Extended physics-informed neural networks (XPINNs): A generalized space-time domain decomposition based deep learning framework for nonlinear partial differential equations. *Commun. Comput. Phys.* 28, 2002-2041 (2020).

- [29] Kharazmi, E., Zhang, Z. & Karniadakis, G. E. hp-VPINNs: Variational physics-informed neural networks with domain decomposition. *Comput. Methods Appl. Mech. Eng.* 374, 113547 (2021).
- [30] Gao, H., Sun, L. & Wang, J. X. PhyGeoNet: Physics-informed geometry-adaptive convolutional neural networks for solving parameterized steady-state PDEs on irregular domain. *J. Comput. Phys.* 428, 110079 (2021).
- [31] Ren, P., Rao, C., Liu, Y., Wang, J. X. & Sun, H. PhyCRNet: Physics-informed convolutional-recurrent network for solving spatiotemporal PDEs. *Comput. Methods Appl. Mech. Eng.* 389, 114399 (2022).
- [32] Kovachki, N. et al. Neural operator: Learning maps between function spaces with applications to pdes. *J. Mach. Learn. Res.* 24, 1-97 (2023).
- [33] Lu, L., Jin, P., Pang, G., Zhang, Z. & Karniadakis, G. E. Learning nonlinear operators via DeepONet based on the universal approximation theorem of operators. *Nat. Mach. Intell.* 3, 218-229 (2021).
- [34] Li, Z. et al. Fourier neural operator for parametric partial differential equations. Preprint at <https://arxiv.org/abs/2010.08895> (2020).
- [35] He, J. et al. Sequential deep operator networks (s-deeponet) for predicting full-field solutions under time-dependent loads. *Eng. Appl. Artif. Intell.* 127, 107258 (2024).
- [36] He, J., Koric, S., Abueidda, D., Najafi, A. & Jasiuk, I. Geom-deeponet: A point-cloud-based deep operator network for field predictions on 3d parameterized geometries. *Comput. Methods Appl. Mech. Eng.* 429, 117130 (2024).
- [37] Peyvan, A., Kumar, V. & Karniadakis, G. E. Fusion-DeepONet: A Data-Efficient Neural Operator for Geometry-Dependent Hypersonic and Supersonic Flows. Preprint at <https://arxiv.org/abs/2501.01934> (2025).
- [38] Sojitra, A., Dhingra, M. & San, O. FEDONet: Fourier-embedded DeepONet for spectrally accurate operator learning. Preprint at <https://arxiv.org/abs/2509.12344> (2025).
- [39] Wang, S., Wang, H. & Perdikaris, P. Learning the solution operator of parametric partial differential equations with physics-informed DeepONets. *Sci. Adv.* 7, eabi8605 (2021).
- [40] Li, Z., Huang, D. Z., Liu, B. & Anandkumar, A. Fourier neural operator with learned deformations for pdes on general geometries. *J. Mach. Learn. Res.* 24, 1-26 (2023).
- [41] Vaswani, A. et al. Attention is all you need. *Adv. Neural Inf. Process. Syst.* 30, 5998-6008 (2017).
- [42] Wang, J., Chakraborty, R. & Stella, X. Y. Transformer for 3D point clouds. *IEEE Trans. Pattern Anal. Mach. Intell.* 44, 4419-4431 (2021).
- [43] Fan, H., Yang, Y. & Kankanhalli, M. Point spatio-temporal transformer networks for point cloud video modeling. *IEEE Trans. Pattern Anal. Mach. Intell.* 45, 2181-2192 (2022).
- [44] Wei, Y., Liu, H., Xie, T., Ke, Q. & Guo, Y. Spatial-temporal transformer for 3d point cloud sequences. In *Proc. IEEE/CVF Winter Conf. Appl. Comput. Vis.* 1171-1180 (2022).
- [45] Hu, X., Zhang, J., Yan, K., Wan, T. & Zheng, X. Physics-Informed Transformer for Efficient Fluid Dynamics Predictions. In *Int. Conf. Wireless Artif. Intell. Comput. Syst. Appl.* 356-368 (Springer, 2025).
- [46] Xu, B., Zhou, Y. & Bian, X. Self-supervised learning based on transformer for flow reconstruction and prediction. *Phys. Fluids* 36, 025111 (2024).
- [47] Luo, Y., Chen, Y. & Zhang, Z. CFDBench: A Large-Scale Benchmark for Machine Learning Methods in Fluid Dynamics. Preprint at <https://arxiv.org/abs/2310.05963> (2023).
- [48] Tali, R. et al. Flowbench: A large scale benchmark for flow simulation over complex geometries. Preprint at <https://arxiv.org/abs/2409.18032> (2024).
- [49] McGreivy, N. & Hakim, A. Weak baselines and reporting biases lead to overoptimism in machine learning for fluid-related partial differential equations. *Nat. Mach. Intell.* 6, 1256 – 1269 (2024).
- [50] Brandstetter, J. Envisioning better benchmarks for machine learning PDE solvers. *Nat. Mach. Intell.* 7, 2-3 (2025).

- [51] Secomb, T. W. Hemodynamics. *Compr. Physiol.* 6, 975-1003 (2016).
- [52] Ku, D. N. Blood flow in arteries. *Annu. Rev. Fluid Mech.* 29, 399-434 (1997).
- [53] Xiang, J. et al. Hemodynamic-morphologic discriminants for intracranial aneurysm rupture. *Stroke* 42, 144 – 152 (2011).
- [54] Rahaman, N. et al. On the spectral bias of neural networks. *Proc. 36th Int. Conf. Mach. Learn.* 97, 5301 – 5310 (2019).
- [55] Wang, S., Wang, H. & Perdikaris, P. On the eigenvector bias of Fourier feature networks: from regression to solving multi-scale PDEs with physics-informed neural networks. *Comput. Methods Appl. Mech. Eng.* 384, 113938 (2021).
- [56] Mathieu, M., Couprie, C. & LeCun, Y. Deep multi-scale video prediction beyond mean square error. *Proc. 4th Int. Conf. Learn. Represent.* (2016).
- [57] Li, Z. et al. Learning spatiotemporal dynamics with a pretrained generative model. *Nat. Mach. Intell.* 6, 1566-1579 (2024).
- [58] Li, X. et al. Aneumo: a large-scale multimodal aneurysm dataset with computational fluid dynamics simulations and deep learning benchmarks. Preprint at arXiv [eess.IV] (2025).
- [59] Ferdian, E., Dubowitz, D. J., Mauger, C. A., Wang, A. & Young, A. A. WSSNet: aortic wall shear stress estimation using deep learning on 4D flow MRI. *Front. Cardiovasc. Med.* 8 (2022).
- [60] Ferdian, E. et al. 4DFlowNet: super-resolution 4D flow MRI using deep learning and computational fluid dynamics. *Front. Phys.* 8 (2020).
- [61] Zimmermann, J. et al. On the impact of vessel wall stiffness on quantitative flow dynamics in a synthetic model of the thoracic aorta. *Sci. Rep.* 11, 6703 (2021).
- [62] Lienen, M., Lüdke, D., Hansen-Palmus, J. & Günnemann, S. From zero to turbulence: generative modeling for 3D flow simulation. *In Int. Conf. Learn. Represent.* (2024).
- [63] Kontogiannis, A., Elgersma, S., Sederman, A. & Juniper, M. 3D Flow MRI data of steady flow through a 3D-printed aorta at Re = 554 and 1526 with SNR 5. *Dataset* (2024).

Acknowledgement

Funding:

This work was supported in part by the the National Natural Science Foundation of China under Grant U2241210, in part by the Basic Research Project of Shenzhen under Grant JCYJ20220818101216034.

Author contributions: F.M. conceived the study, designed the experiments, and drafted the manuscript. C.C. and Yu.L. conducted model development and coding. Y.Z. and J.C. revised the manuscript. Y.H. and J.H. conducted the CFD simulations. H.Z. collected the data. J.Y. provided funding and co-supervised the research. H.L. reviewed and provided critical revision of the manuscript. Ye.L. supervised the research. All authors reviewed and approved the final manuscript.

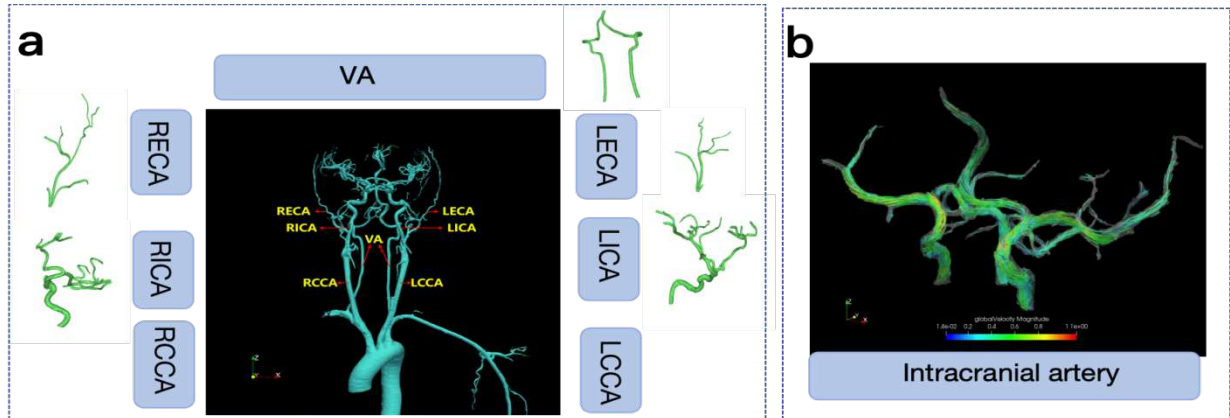
Competing Interests:

The authors declare that they have no competing interests.

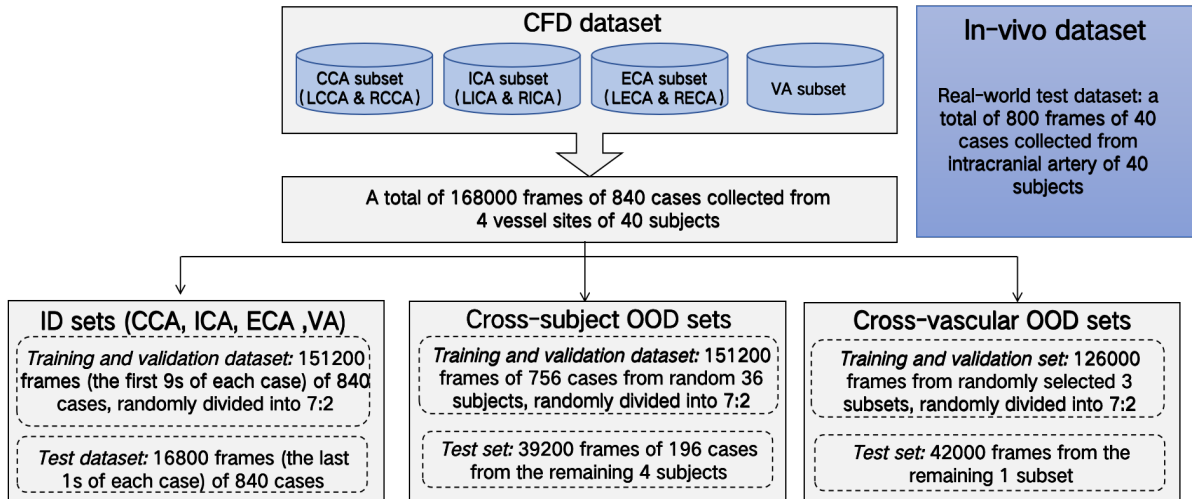
Data and material availability:

Data availability. The main data supporting the results in this study are available within the paper and its Supplementary Information. Our 4DHemoDB database for training, validating, and testing the model is available on Hugging Face upon request (<https://huggingface.co/datasets/PhysAI/4DHemoDB>). We have deposited a minimum dataset for model validation on GitHub (<https://github.com/4DHemoAI/4DHemoX>).

Code availability. The source code to train 4DHemoFormer and baselines and reproduce the results is available in GitHub at <https://github.com/4DHemoAI/4DHemoX>.



Extended Data Fig. 1. | a, Detailed description of vessel sites in the CFD dataset. b, Vessel sites for In-vivo dataset.



Extended Data Fig. 2. | Flow chart describing the datasets collected for the development and evaluation of the model. ID, in-distribution; OOD, out-of-distribution; ICA, internal carotid artery; ECA, external carotid artery; VA, vertebral artery; CCA, common carotid artery.

Extended Data Table 1. Comparison of 4DHemoDB and existing dataset

Name	Year	Organization	Dimensions	Types	Cases	Frames	TR	SR
CFD Bench ⁴⁷	2023	Tsing Hua University	2D	Idealized CFD simulations	738	690800	-	64 ² uniform grid
Flow Bench ⁴⁸	2024	Iowa State University	2D,3D	Idealized CFD simulations	10650	-	-	128 ² ,256 ² ,512 ² uniform grid
Aneumo ⁵⁸	2025	Fudan University	3D	Patient-specific CFD simulations	427	-	-	0.15mm minimum grid
WSSNet ⁵⁹	2021	University of Auckland	3D	Patient-specific CFD simulations	43	3096	40ms	1.2mm,2.4 mm
4DFlowNet ⁶⁰	2020	University of Auckland	3D	Patient-specific CFD simulations	11	792	10ms	2mm,4mm
AS4DF ⁶¹	2021	Stanford University	3D	Patient-specific CFD simulations	3	-	20ms/40ms/62.5ms	2.5mm
3D Turbulent Flow Simulations ⁶²	2024	Technical University of Munich	3D	Idealized CFD simulations	45	225000	0.1ms	-
3D flow MRI data ⁶³	2024	University of Cambridge	3D	In-vivo 4D flow MRI data	2	6	3.46h	0.1mm
4DHemoDB	2026	UESTC	3D	Patient-specific CFD simulations and In-vivo 4D flow MRI data	840	168000	50ms	CFD: 0.45mm maximum grid in vivo data: 0.6mm

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Supplementaryinformation4DHemoX.pdf](#)