TumorFlow: Physics-Guided Longitudinal MRI Synthesis of Glioblastoma Growth

Valentin Biller[1]*, Niklas Bubeck[2,4]*, Lucas Zimmer[3], Ayhan Can Erdur[4], Sandeep Nagar[3], Anke Meyer-Baese[6], Daniel Rueckert[2,4,5], Benedikt Wiestler[2,3]**, and Jonas Weidner[2,3]**

[1] Technical University of Munich (TUM), Germany
[2] Munich Center for Machine Learning (MCML)
[3] AI for Image-Guided Diagnosis and Therapy, TUM, Germany
[4] Chair for AI in Healthcare and Medicine, TUM and TUM University Hospital, Munich, Germany
[5] Imperial College London
[6] Florida State University
* equal contribution
** equal contribution
{valentin.biller, niklas.bubeck, j.weidner}@tum.de

Abstract. Glioblastoma exhibits diverse, infiltrative, and patient-specific growth patterns that are only partially visible on routine MRI, making it difficult to reliably assess true tumor extent and personalize treatment planning and follow-up. We present a biophysically-conditioned generative framework that synthesizes biologically realistic 3D brain MRI volumes from estimated, spatially continuous tumor-concentration fields. Our approach combines a generative model with tumor-infiltration maps that can be propagated through time using a biophysical growth model, enabling fine-grained control over tumor shape and growth while preserving patient anatomy. This enables us to synthesize consistent tumor growth trajectories directly in the space of real patients, providing interpretable, controllable estimation of tumor infiltration and progression beyond what is explicitly observed in imaging. We evaluate the framework on longitudinal glioblastoma cases and demonstrate that it can generate temporally coherent sequences with plausible changes in tumor appearance and surrounding tissue response. These results suggest that integrating mechanistic tumor growth priors with modern generative modeling can provide a tool for generating controlled synthetic data to support downstream neuro-oncology workflows. In longitudinal extrapolation, we achieve a consistent 75% Dice overlap with the biophysical model while maintaining a constant PSNR of 25 in the surrounding tissue. Our code is available at: https://github.com/valentin-biller/lgm.git

Keywords: Glioblastoma · Tumor Growth Modeling · Longitudinal Brain MRI Synthesis · Biophysical Conditioning

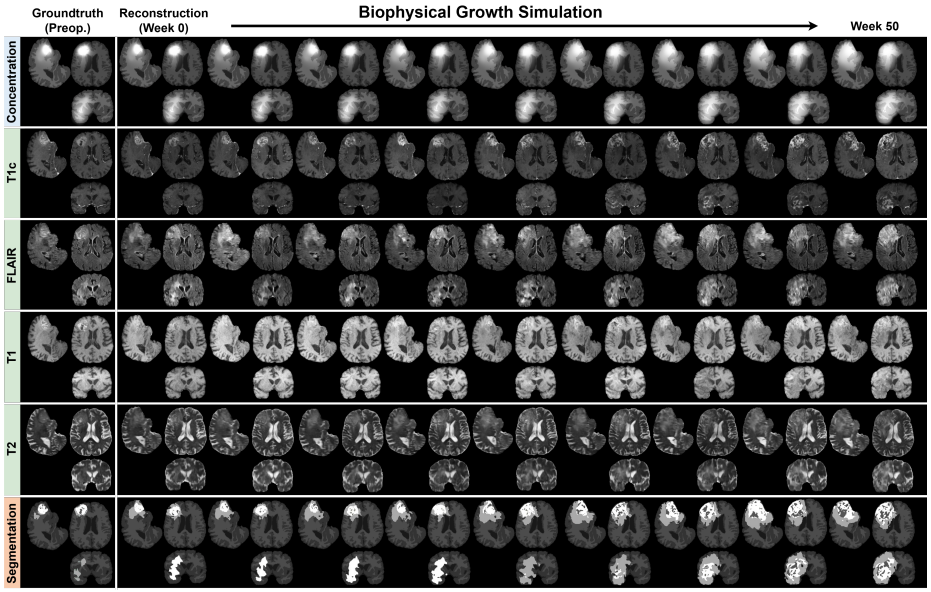

Fig. 1: **Representative Longitudinal Tumor Growth Prediction**. Starting from the preoperative state (week 0), a biophysical growth model simulates time-resolved tumor concentration fields (blue), serving as conditioning for image synthesis. The model generates T1c, FLAIR, T1, T2 modalities (green), preserving anatomy while capturing modality-specific progression. A standard segmentation model applied to the generated images yields time-consistent masks (orange).

1 Introduction

Glioblastoma is characterized by a highly infiltrative growth pattern. While MRI is essential for diagnosis, treatment planning, and longitudinal monitoring, it fails to detect the true extent and spread of glioblastoma and is particularly blind to the diffuse invasion of tumor cells into the surrounding brain [25].

Biophysical brain tumor modeling has emerged as an interpretable and robust approach for estimating patient-specific tumor cell distributions and growth trajectories, enabling counterfactual predictions of tumor progression [13,3,23]. These models, however, typically operate on a continuous tumor-concentration field, not in the imaging space used in clinical practice for decision-making.

Recently, generative models have emerged for brain tumor MRI [21,5,18,9]. However, they are conditioned on tumor segmentation masks, which constrain control to coarse morphology and prevent modeling spatially continuous infiltration patterns or their smooth evolution over time, a key requirement for *in silico* modeling of tumor response and personalized treatment decisions. In parallel, methods that are conditioned only on tumor volume were developed [24,10,12]. These approaches aim to implicitly infer biophysical growth characteristics from data, but they often fail to produce temporally coherent progression and lack

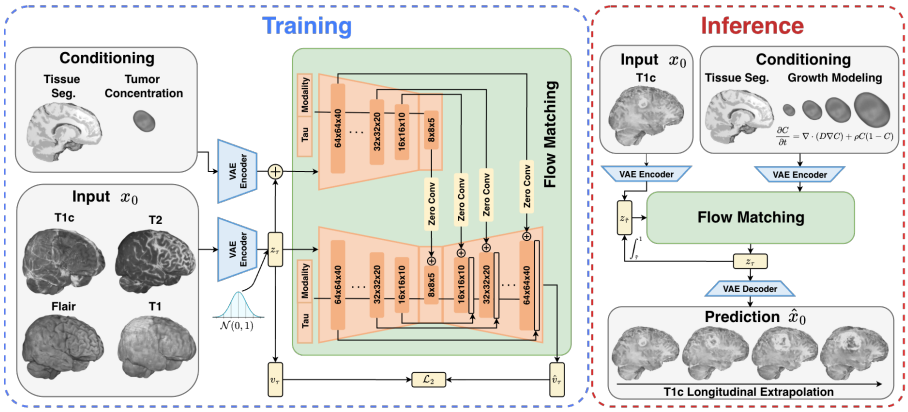

Fig. 2: **Architecture of TumorFlow.** Modality, tissue segmentations, and tumor concentrations are provided as conditioning inputs to the latent generative model. Orange boxes denote trainable modules, whereas blue boxes indicate frozen, pre-trained components. During inference, a full growth trajectory can be extrapolated from a biophysical simulation.

modifiability (e.g., for simulating different treatments). This limitation is further exacerbated by the scarcity and heterogeneity of longitudinal datasets.

Our main contributions are: **(i)** We train a generative flow matching model for 3D glioblastoma MRI on preoperative volumes only, which outperforms state-of-the-art approaches on distributional image quality and conditioning adherence metrics. **(ii)** We condition the model on a tumor infiltration field, enabling fine-grained control over tumor generation significantly beyond what classical segmentation masks offer. **(iii)** We combine the generative model with state-of-the-art biophysical brain tumor modeling to generate biologically plausible longitudinal growth trajectories from cross-sectional MRI alone.

2 Method

We present a two-stage 3D latent generative framework (Fig. 2) for brain tumor MRI synthesis and progression modeling. A pretrained 3D VAE encodes multimodal MR volumes, and a latent flow matching model conditions on modality, tissue segmentation, and biophysically-derived tumor-concentration fields temporally propagated forward during inference for plausible growth prediction.

Autoencoding. Let $\mathbf{x} \in \mathbb{R}^{256\times256\times160}$ denote a single-modality 3D MRI volume. We utilize a pretrained variational autoencoder (VAE) from the MAISI framework [8] which maps $\mathbf{x}$ into latent space $\mathbf{z}$ given the encoder $\mathcal{E}_\phi$ approximating the posterior distribution $q_\phi(z|x)$ with $z \sim \mathcal{N}(\mu_z, \Sigma_z)$ from which latent codes $\mathbf{z} \in \mathbb{R}^{4\times64\times64\times40}$ are sampled via the reparameterization trick. Afterwards,

the image can be reconstructed again by estimating the likelihood $p_\theta(x|z)$ using the decoder $\mathcal{D}_\theta$, where ϕ and θ are learnable parameters.

Latent Flow Matching. For generation, we utilize an Optimal Transport Flow Matching (OT-FM) formulation. Let $\mathbf{z}_1$ be a real MRI latent and $\mathbf{z}_0 \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ a Gaussian prior. We define an affine interpolation between source and target as $\mathbf{z}_\tau = (1-\tau)\mathbf{z}_0 + \tau\mathbf{z}_1$ with $\tau \in [0,1]$ and ground-truth velocity $\frac{d\mathbf{z}_\tau}{d\tau} = \mathbf{z}_1 - \mathbf{z}_0$. A neural network $v_\psi(\mathbf{z}_\tau, \tau, \mathbf{c})$ with learnable parameters ψ is conditioned on $\mathbf{c}$ to predict the velocity field such that training minimizes

$$\mathcal{L}_{\mathrm{FM}}(\psi) = \mathbb{E}_{\mathbf{z}_0,\mathbf{z}_1,\tau,\mathbf{c}}\left[\left\|v_\psi(\mathbf{z}_\tau,\tau,\mathbf{c}) - (\mathbf{z}_1 - \mathbf{z}_0)\right\|_2^2\right], \quad \tau \sim \mathcal{U}[0,1]. \tag{1}$$

At inference, a new sample $\hat{\mathbf{z}}_1$ can be generated by drawing from source and integrating over the learned velocity field. $\hat{\mathbf{z}}_1 = \mathbf{z}_0 + \int_0^1 v_\psi(\mathbf{z}_\tau, \tau, \mathbf{c})\, d\tau$. The resulting latent $\hat{\mathbf{z}}_1$ is decoded by the VAE to reconstruct the final volume.

Conditioning. The conditioning input $\mathbf{c} = (m, \mathbf{c}_s, \mathbf{c}_{\mathrm{tc}})$ comprises MRI modality m, tissue segmentation $\mathbf{c}_s$, and tumor concentration map $\mathbf{c}_{\mathrm{tc}}$. The learnable modality embedding $\mathbf{e}_m = \mathrm{Embed}(m)$ is concatenated with the timestep embedding. Both $\mathbf{c}_s$ (from atlas registration [1]) and $\mathbf{c}_{\mathrm{tc}}$ share the pretrained encoder $\mathcal{E}_\phi$ used to encode the MRI volumes for representation consistency, yielding $\mathbf{z}_s = \mathcal{E}_\phi(\mathbf{c}_s)$ and $\mathbf{z}_{\mathrm{tc}} = \mathcal{E}_\phi(\mathbf{c}_{\mathrm{tc}})$. Their concatenation forms the spatial conditioning tensor $\mathbf{c}_{\mathrm{spatial}} = \mathrm{concat}(\mathbf{z}_s, \mathbf{z}_{\mathrm{tc}}) \in \mathbb{R}^{8\times64\times64\times40}$, which is injected into the U-Net backbone via a ControlNet-style branch [26]: $\mathbf{h}_\ell \leftarrow \mathbf{h}_\ell + \mathcal{F}_\ell(\mathbf{c}_{\mathrm{spatial}})$, where $\mathcal{F}_\ell$ are learnable projections at scale ℓ.

Biophysical Tumor Growth Modeling. We model glioblastoma growth using the Fisher-Kolmogorov reaction-diffusion equation, which is widely adopted in the field [7,3,13]. It models the temporal change of the tumor concentration C based on a reaction and a diffusion term:

$$\frac{\partial C}{\partial t} \;=\; \nabla \cdot (D\,\nabla C) \;+\; \rho\, C\,(1-C), \tag{2}$$

where D is a tissue-dependent diffusion coefficient, and ρ the proliferation rate.

Longitudinal Image Generation. We must balance the temporal consistency of patient anatomy with the growing tumor. Let $\{\mathbf{c}_{\mathrm{tc}}^t\}_{t=0}^T$ denote time-indexed t tumor concentration fields. The initial concentration $\mathbf{c}_{\mathrm{tc}}^0$ is estimated by a biophysical growth model [22] to match the pre-op tumor segmentation. Each $\mathbf{c}_{\mathrm{tc}}^t$ is encoded into latent space and used for conditional synthesis. For temporal coherence and synthesis of time $t+1$ the previous state z_1^t gets transported backward along the deterministic flow trajectory into a corrupted version $z_{\tilde{\tau}}^t$. Given this intermediate representation, we integrate forward under the updated conditioning, enabling controlled temporal evolution while preserving anatomical consistency:

$$\hat{\mathbf{z}}_1^{t+1} = \mathbf{z}_{\tilde{\tau}}^t + \int_{\tilde{\tau}}^1 v_\psi(\mathbf{z}_\tau^t, \tau, \mathbf{c}_{tc}^t)\, d\tau. \tag{3}$$

Recursive application yields anatomically and temporally consistent longitudinal synthesis, with corruption level $\tilde{\tau}$ selected empirically for each method.

Table 1: **Quantitative Evaluation of Generated Volumes.** Tumor Dice compares thresholded tumor concentration fields with pretrained BraTS segmentations. Tissue Dice (CSF, GM, WM) compares the conditioning segmentations against segmentations re-derived via atlas registration. FID, KID, and MS-SSIM assess distributional alignment, perceptual similarity, and diversity against the test set. The reference metrics are empirical upper bounds derived from test-set calibration experiments, reflecting limits by VAE reconstruction fidelity, uncertainty in tumor-concentration derivations, and automatic tumor segmentation.

Model	**Dice ↑**				**FID ↓**	**KID ↓**	**MS-SSIM ↑**
	Tumor	**CSF**	**GM**	**WM**			
Reference	0.779	–	–	–	0.330	0.001	–
TumorFlow	**0.739**	0.648	0.802	0.824	**2.125**	**0.020**	**0.675**
Med-DDPM [5]	0.344	0.310	0.653	0.659	106.246	1.618	0.588
OT-FM w. ViT	0.572	0.648	0.802	**0.829**	2.540	0.026	0.675
DDPM w. U-Net	0.579	**0.655**	**0.805**	0.828	11.221	0.083	0.648

3 Experiments

Implementation Details. We use the publicly available BraTS 2021 dataset [2] as well as UCSF-PDGM [4], TCGA-GBM [16], TCGA-LGG [14], and Rembrandt [15], excluding overlapping patients. This increases distributional diversity [20] and mitigates memorization risk [19]. The combined dataset comprises 3,602 subjects, split into 80% training (2,881) and 20% validation (721). All volumes are co-registered to a common template, resampled to $1\,\mathrm{mm}^3$ isotropic resolution, and skull-stripped. We include Lumiere [17] for longitudinal evaluation. Models are trained on a single NVIDIA A100 (40 GB) using AdamW (learning rate 10^{-4}, zero weight decay) with cosine annealing and an EMA decay of $\gamma = 0.999$ for stability. U-Net-based architectures contain 236M parameters, the ViT-based variant 145M, and Med-DDPM 58M. Flow matching models converge within approximately one week; the DDPM variant requires three weeks.

References, Baseline and Variations. The held-out test set serves as a calibration baseline. Unless stated otherwise, Fréchet Inception Distance (FID) and Kernel Inception Distance (KID) are reported in $\times 10^{-3}$. Pretrained VAE reconstructions yield PSNR 32.09 and FID/KID 0.330/0.001, establishing the empirical fidelity ceiling. For tumor modeling, Dice overlap between reconstructed concentration fields and MR-based automatic segmentations [11] reaches 0.779, reflecting the uncertainty-imposed upper bound on segmentation accuracy.

We compare against the BraTS-pretrained Med-DDPM baseline [5] by thresholding the tumor concentration map to obtain a segmentation mask for conditioning (edema: $[0.2, 0.6)$, enhancing tumor: ≥ 0.6). These thresholds are fixed from the biophysical concentration map to not disadvantage the baseline.

We further compare against two architectural variants: 1. switching the corruption paradigm to a Denoising Diffusion Probabilistic Model (DDPM), and

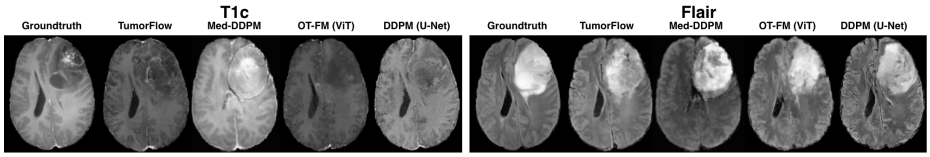

Fig. 3: **Qualitative Comparison**. Synthetic volumes generated by all four models from Table 1, conditioned on the same $\mathbf{c}_s$ and $\mathbf{c}_{\text{tc}}$ inputs.

2. replacing the 3D U-Net backbone with a Vision Transformer (ViT) [6] using its base configuration and patch size of 2. The limited number of SOTA baselines reflects the absence of publicly available pretrained weights for 3D spatially conditional medical generators; the DDPM and ViT variants were therefore implemented as principled architectural comparisons.

Generation of Static Brain Tumor MRI. First, we assess the general *generative performance* of the model by evaluating image fidelity and diversity on the static, non-longitudinal setup. We compute FID, KID, and MS-SSIM between the generated volumes and the real test volumes to quantify distributional fidelity and diversity of synthesized samples. Second, we evaluate *adherence* to the conditioning signals and compare tumor volumes from the input concentration field and from the generated images. The reference mask is obtained by thresholding the concentration field at 0.2, which is the calibration threshold for the biophysical growth model; by using the same thresholds, all methods are evaluated against the same target masks, ensuring fairness. Tumor volumes in the synthesized images are estimated with a pretrained segmentation model from BraTS Orchestrator [11], using the top-performing method from the Adult Glioma Segmentation Challenge [2]. Tumor adherence is quantified via the Dice Score [27] between reference mask and predicted segmentation. Tissue-specific Dice (CSF, GM, WM) between the conditioning tissue segmentations and those derived from the synthesized images quantifies healthy tissue adherence.

Quantitative evaluation of fidelity, diversity, and conditioning adherence shows that our model outperforms the Med-DDPM [5] baseline as well as the diffusion and ViT variations in brain MRI synthesis (Table 1). Lower FID and KID scores indicate a closer match to real data, while higher MS-SSIM values confirm improved perceptual and structural fidelity. Tissue-specific Dice further confirms *healthy tissue adherence*, demonstrating preservation of healthy anatomy through tissue-segmentation conditioning. Additionally, our model demonstrates substantially improved *tumor adherence* compared to the Med-DDPM [5] baseline, with generated tumor regions aligning more closely to the intended geometry, as shown qualitatively in Fig. 3. Leveraging tumor concentration maps yields higher Dice Score between the target signal and the synthesized tumor, reflecting tighter spatial agreement. Even accounting for potential artifacts from resampling Med-DDPM's inference resolution ($192 \times 192 \times 144$) to our training resolution ($240 \times 240 \times 155$), TumorFlow exceeds the original Med-DDPM paper's reported Dice of 0.62 ± 0.29. Crucially, the volumes in Fig. 3 are not expected to reconstruct

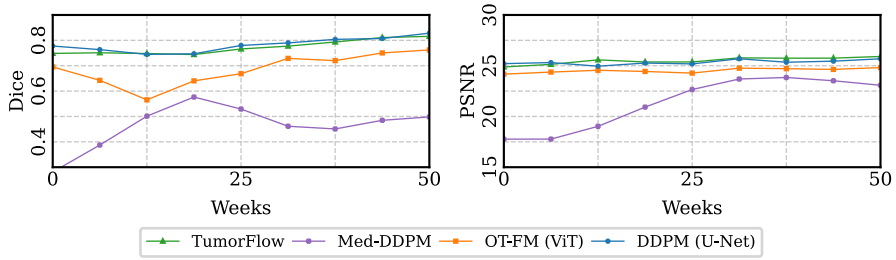

Fig. 4: **Left: Tumor Consistency.** Dice over time, assessing tumor growth consistency. **Right: Tissue Consistency.** PSNR over time, measuring structural drift in healthy tissue. Stable values indicate coherent generation.

the ground truth; predicting MRI from concentration fields is inherently ill-posed, as tumor appearance varies across biology, acquisition, and treatment. Our goal is controllable generation that adheres to a tumor growth field while preserving anatomy rather than voxel-identical reconstruction.

Longitudinal Extrapolation. The TumorGrowthToolkit [3] is configured with $\rho = 0.03\,\mathrm{d}^{-1}$ and $D = 0.28\,\mathrm{mm}^2\mathrm{d}^{-1}$ to generate physics-informed spatiotemporal tumor concentration fields $c^t(\mathbf{x})_{t=0}^{T}$ for synthesizing longitudinal MRI sequences. Fixed proliferation and diffusion parameters enable controlled synthesis and isolate the generative model's contribution. *Temporal consistency* is enforced via predecessor-manipulation, where each time point is conditioned on the evolved concentration field and initialized at $\tilde{\tau} = 0.15$.

In Fig. 1, we show a representative longitudinal prediction over 50 weeks. The biophysical simulation produces a smooth tumor cell concentration trajectory that informs the generation of follow-up scans. Despite substantial changes in the extent of the tumor, the patient anatomy outside the tumor remains stable over time. The growing tumor, in particular the necrotic and contrast-enhancing areas, exhibits realistic morphology and imaging characteristics. The downstream segmentations provide an additional consistency check, producing masks that evolve plausibly and align well with the predicted growth pattern.

In the left panel of Fig. 4, our model achieves the highest Dice, demonstrating superior *temporal consistency* and indicating that the evolving, synthetic tumors are highly plausible. The right panel quantifies temporal tissue consistency in non-tumorous regions between consecutive time points using PSNR. Excessively high PSNR values imply an overly static anatomy that hinders realistic tumor growth, while low values reflect excessive structural drift and poor subject coherence. Our model achieves a balanced trade-off, maintaining stable Dice and PSNR curves that indicate temporally coherent synthesis. Over time, consistent metric values suggest controlled tumor progression and preserved anatomy, whereas deviations reveal either insufficient tumor growth dynamics or instability in structural tissue fidelity. To verify this spatial precision, we analyzed distant

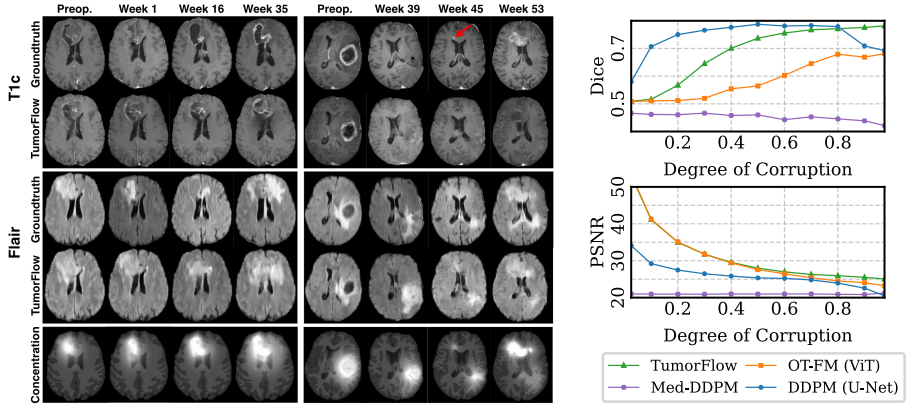

Fig. 5: **Left: Real Patients.** Comparing TumorFlow and ground truth tumor development in two patients from Lumiere [17]. The red arrow points to a new tumor lesion in the corpus callosum. **Right: Effect of Corruption Initialization on Longitudinal Generation.** Top: Dice between generated tumors and conditioning. Bottom: PSNR between consecutive non-tumorous regions.

healthy tissue separately from tumor-adjacent tissue. TumorFlow yields a higher PSNR in distant tissue than in adjacent tissue (18.66 vs. 17.74), consistent with expected local changes near simulated tumor growth and infiltration, and broad preservation elsewhere. Notably, this spatial trend is shared by the architectural variations but is absent in Med-DDPM.

As treatment-naive longitudinal public data do not exist, we evaluate our method on the postoperative Lumiere Dataset (see Fig. 5 left). Tumor concentration is estimated by independently fitting a biophysical model at each time step, with generation initialized from the pre-op image. Despite the resection cavity and post-surgery deformation - unseen during training - we observe temporally consistent and anatomically plausible longitudinal images. In the right patient, a satellite lesion in the corpus callosum at week 45 (red arrow) closely follows the predicted concentration growth pattern while surrounding anatomy remains visually consistent. For quantitative validation, we compute Dice scores between expert Lumiere annotations and tumor volumes segmented from the synthesized images using the pretrained BraTS model. Despite the strong domain shift (unseen post-op cavities), TumorFlow achieves Dice 0.42 vs. 0.22 for Med-DDPM, and 0.78 vs. 0.36 preoperatively. These results confirm strong conditioning responsiveness, despite spatial alignment errors from the Fisher-Kolmogorov model and the remaining challenge of post-operative progression.

Ablation. In the optimal transport corruption process, fine-grained features degrade earlier than coarse structures. We assess this by varying the corruption level $\tilde{\tau}$ used to initialize the predecessor manipulation (Fig. 5 right). Low $\tilde{\tau}$ values yield volumes nearly identical to the previous state (high PSNR, minimal tumor

progression), while higher values induce stronger, more realistic growth dynamics. TumorFlow exhibits a smooth, monotonic Dice response across corruption levels, indicating stable and controllable tumor evolution. In contrast, DDPM-based models show irregular trends reflecting a weaker balance between anatomical stability and tumor fidelity, with Med-DDPM [5] diverging entirely.

4 Conclusion

We introduced a physics-guided generative framework for longitudinal glioblastoma MRI synthesis that conditions high-fidelity 3D image generation on spatially continuous tumor concentration fields and evolves these fields over time with a Fisher-Kolmogorov reaction-diffusion growth model. By conditioning on spatially continuous tumor concentration fields and evolving them with a biophysical growth prior, we can generate follow-up images that reflect biologically plausible tumor growth trajectories while keeping the individual anatomy and image appearance stable over time. Importantly, we achieve this while training only on cross-sectional, preoperative data, effectively alleviating the persistent shortage of high-quality longitudinal datasets in medical imaging.

While our framework shows the potential of coupling biophysical models with conditional generation, clinical translation requires replacing fixed parameters with patient-specific fitting. Future work will integrate treatment effects, resection-aware concentrations, and post-operative training scans. Ultimately, our work lays an important foundation for *in silico* disease and treatment modeling.

Acknowledgments. This work is funded in part by the European Research Council (ERC) project Deep4MI (884622) and the Munich Center for Machine Learning.

Disclosure of Interests. The authors declare no competing interests.

References

1. Avants, B.B., Epstein, C.L., Grossman, M., Gee, J.C.: Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. Medical image analysis **12**(1), 26–41 (2008)
2. Baid, U., Ghodasara, S., Mohan, S., Bilello, M., Calabrese, E., Colak, E., Farahani, K., Kalpathy-Cramer, J., Kitamura, F.C., Pati, S., et al.: The rsna-asnr-miccai brats 2021 benchmark on brain tumor segmentation and radiogenomic classification. arXiv preprint arXiv:2107.02314 (2021)
3. Balcerak, M., Weidner, J., Karnakov, P., Ezhov, I., Litvinov, S., Koumoutsakos, P., Amiranashvili, T., Zhang, R.Z., Lowengrub, J.S., Yakushev, I., et al.: Individualizing glioma radiotherapy planning by optimization of a data and physics-informed discrete loss. Nature Communications **16**(1), 5982 (2025)
4. Calabrese, E., Villanueva-Meyer, J.E., Rudie, J.D., Rauschecker, A.M., Baid, U., Bakas, S., Cha, S., Mongan, J.T., Hess, C.P.: The university of california san francisco preoperative diffuse glioma mri dataset. Radiology: Artificial Intelligence **4**(6), e220058 (2022)
5. Dorjsembe, Z., Pao, H.K., Odonchimed, S., Xiao, F.: Conditional diffusion models for semantic 3d brain mri synthesis. IEEE Journal of Biomedical and Health Informatics **28**(7), 4084–4093 (2024)
6. Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., et al.: An image is worth 16x16 words: Transformers for image recognition at scale. arXiv preprint arXiv:2010.11929 (2020)
7. Ezhov, I., Scibilia, K., Franitza, K., Steinbauer, F., Shit, S., Zimmer, L., Lipkova, J., Kofler, F., Paetzold, J.C., Canalini, L., et al.: Learn-morph-infer: a new way of solving the inverse problem for brain tumor modeling. Medical Image Analysis **83**, 102672 (2023)
8. Guo, P., Zhao, C., Yang, D., Xu, Z., Nath, V., Tang, Y., Simon, B., Belue, M., Harmon, S., Turkbey, B., et al.: Maisi: Medical ai for synthetic imaging. In: 2025 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV). pp. 4430–4441. IEEE (2025)
9. Jiang, L., Mao, Y., Wang, X., Chen, X., Li, C.: Cola-diff: Conditional latent diffusion model for multi-modal mri synthesis. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 398–408. Springer (2023)
10. Kebaili, A., Lapuyade-Lahorgue, J., Vera, P., Ruan, S.: 3d mri synthesis with slice-based latent diffusion models: Improving tumor segmentation tasks in data-scarce regimes. In: 2024 IEEE International Symposium on Biomedical Imaging (ISBI). pp. 1–5. IEEE (2024)
11. Kofler, F., Rosier, M., Astaraki, M., Baid, U., Möller, H., Buchner, J.A., Steinbauer, F., Oswald, E., de la Rosa, E., Ezhov, I., et al.: Brats orchestrator: Democratizing and disseminating state-of-the-art brain tumor image analysis. arXiv preprint arXiv:2506.13807 (2025)
12. Laslo, D., Georgiou, E., Linguraru, M.G., Rauschecker, A., Muller, S., Jutzeler, C.R., Bruningk, S.: Mechanistic learning with guided diffusion models to predict spatio-temporal brain tumor growth. arXiv preprint arXiv:2509.09610 (2025)
13. Lipkova, J., Angelikopoulos, P., Wu, S., Alberts, E., Wiestler, B., Diehl, C., Preibisch, C., Pyka, T., Combs, S.E., Hadjidoukas, P., et al.: Personalized radiotherapy design for glioblastoma: integrating mathematical tumor models, multimodal scans, and bayesian inference. IEEE transactions on medical imaging **38**(8), 1875–1884 (2019)

14. Pedano, N., Flanders, A.E., Scarpace, L., Mikkelsen, T., Eschbacher, J.M., Hermes, B., Sisneros, V., Barnholtz-Sloan, J., Ostrom, Q.: The cancer genome atlas low grade glioma collection (tcga-lgg). (No Title) (2016)
15. Sayah, A., Bencheqroun, C., Bhuvaneshwar, K., Belouali, A., Bakas, S., Sako, C., Davatzikos, C., Alaoui, A., Madhavan, S., Gusev, Y.: Enhancing the rembrandt mri collection with expert segmentation labels and quantitative radiomic features. Scientific Data **9**(1), 338 (2022)
16. Scarpace, L., Mikkelsen, T., Cha, S., Rao, S., Tekchandani, S., Gutman, D., Saltz, J.H., Erickson, B.J., Pedano, N., Flanders, A.E., et al.: The cancer genome atlas glioblastoma multiforme collection (tcga-gbm). The Cancer Imaging Archive (2016)
17. Suter, Y., Knecht, U., Valenzuela, W., Notter, M., Hewer, E., Schucht, P., Wiest, R., Reyes, M.: The lumiere dataset: Longitudinal glioblastoma mri with expert rano evaluation. Scientific data **9**(1), 768 (2022)
18. Truong, N.C., Yogananda, C.G.B., Wagner, B.C., Holcomb, J.M., Reddy, D., Saadat, N., Hatanpaa, K.J., Patel, T.R., Fei, B., Lee, M.D., et al.: Synthesizing 3d multicontrast brain tumor mris using tumor mask conditioning. In: Medical Imaging 2024: Imaging Informatics for Healthcare, Research, and Applications. vol. 12931, pp. 116–120. SPIE (2024)
19. Usman Akbar, M., Larsson, M., Blystad, I., Eklund, A.: Brain tumor segmentation using synthetic mr images-a comparison of gans and diffusion models. Scientific Data **11**(1), 259 (2024)
20. Usman Akbar, M., Wang, W., Eklund, A.: Beware of diffusion models for synthesizing medical images—a comparison with gans in terms of memorizing brain mri and chest x-ray images. Machine Learning: Science and Technology **6**(1), 015022 (2025)
21. Wang, H., Liu, Z., Sun, K., Wang, X., Shen, D., Cui, Z.: 3d meddiffusion: A 3d medical latent diffusion model for controllable and high-quality medical image generation. IEEE Transactions on Medical Imaging (2025)
22. Weidner, J., Balcerak, M., Ezhov, I., Datchev, A., Lux, L., Rueckert, L.Z.D., Menze, B., Wiestler, B.: Spatial brain tumor concentration estimation for individualized radiotherapy planning. arXiv e-prints pp. arXiv–2412 (2024)
23. Weidner, J., Ezhov, I., Balcerak, M., Metz, M.C., Litvinov, S., Kaltenbach, S., Feiner, L., Lux, L., Kofler, F., Lipkova, J., et al.: A learnable prior improves inverse tumor growth modeling. IEEE Transactions on Medical Imaging **44**(3), 1297–1307 (2024)
24. Wolleb, J., Sandkühler, R., Bieder, F., Cattin, P.C.: The swiss army knife for image-to-image translation: Multi-task diffusion models. arXiv preprint arXiv:2204.02641 (2022)
25. Yan, J.L., Li, C., Boonzaier, N.R., Fountain, D.M., Larkin, T.J., Matys, T., van der Hoorn, A., Price, S.J.: Multimodal mri characteristics of the glioblastoma infiltration beyond contrast enhancement. Therapeutic advances in neurological disorders **12**, 1756286419844664 (2019)
26. Zhang, L., Rao, A., Agrawala, M.: Adding conditional control to text-to-image diffusion models. In: Proceedings of the IEEE/CVF international conference on computer vision. pp. 3836–3847 (2023)
27. Zou, K.H., Warfield, S.K., Bharatha, A., Tempany, C.M., Kaus, M.R., Haker, S.J., Wells III, W.M., Jolesz, F.A., Kikinis, R.: Statistical validation of image segmentation quality based on a spatial overlap index1: scientific reports. Academic radiology **11**(2), 178–189 (2004)