Ensemble inversion for brain tumor growth models with mass effect

Shashank Subramanian,
Oden Institute, University of Texas at Austin, Texas, USA

Ali Ghafouri,
Oden Institute, University of Texas at Austin, Texas, USA

Klaudius Scheufele,
Oden Institute, University of Texas at Austin, Texas, USA

Naveen Himthani,
Oden Institute, University of Texas at Austin, Texas, USA

Christos Davatzikos [Fellow, IEEE],
University of Pennsylvania, Philadelphia, USA

George Biros [Member, IEEE]
Oden Institute, University of Texas at Austin, Texas, USA

Abstract

We propose a method for extracting physics-based biomarkers from a *single* multiparametric Magnetic Resonance Imaging (mpMRI) scan bearing a glioma tumor. We account for mass effect, the deformation of brain parenchyma due to the growing tumor, which on its own is an important radiographic feature but its automatic quantification remains an open problem. In particular, we calibrate a partial differential equation (PDE) tumor growth model that captures mass effect, parameterized by *a single scalar parameter*, tumor *proliferation*, *migration*, while localizing the *tumor initiation site*. The single-scan calibration problem is severely ill-posed because the precancerous, healthy, brain anatomy is unknown. To address the ill-posedness, we introduce an ensemble inversion scheme that uses a number of *normal subject brain templates* as proxies for the healthy precancer subject anatomy. We verify our solver on a synthetic dataset and perform a retrospective analysis on a clinical dataset of 216 glioblastoma (GBM) patients. We analyze the reconstructions using our calibrated biophysical model and demonstrate that our solver provides both global and local quantitative measures of tumor biophysics and mass effect. We further highlight the improved performance in model calibration through the inclusion of mass effect in tumor growth models—including mass effect in the model leads to 10% increase in average dice coefficients for patients with significant mass effect. We further evaluate our model by introducing novel biophysics-based features and using them for survival analysis. Our preliminary analysis suggests that including such features can improve patient stratification and survival prediction.

shashanksubramanian@utexas.edu .

Keywords

tumor growth model personalization; inverse problem; mass effect; glioblastoma

---

1. INTRODUCTION

Computational oncology is an emerging field that attempts to integrate biophysical models with imaging with the goal of assisting image analysis in a clinical setting. Typically, the integration is accomplished by calibrating PDE model parameters from images. A significant challenge in brain tumors, and in particular, GBMs, is the single-scan calibration, that becomes even harder in the presence of mass effect. However, modeling provides the capability for automatic quantification of mass effect along with additional biomarkers related to infiltration and tumor aggressiveness. Mass effect alone, is quite significant as it is an important radiographic marker [1]–[3]. Here, we propose a biophysical model calibrated using a single pretreatment scan. Longitudinal pretreatment scans are rare for GBMs since immediate treatment that includes surgical resection and subsequent radiotherapy, is necessary for most patients.

A. Contributions

The single-scan calibration problem is formidable for two main reasons: the tumor initial condition (IC) and the subject's healthy precancerous brain anatomy defined by the brain anatomy/structural MRI before the cancer begins to grow, are unknown. To circumvent this, a template (normal subject brain from another healthy individual) is used as a *proxy* to the healthy subject brain. However, natural anatomical differences between the template and the subject interfere with tumorrelated deformations; disentangling the two is hard. In light of these difficulties, our contributions, as an extension to our recently published work [4]), are as follows:

1. Based on the method described in [5], we propose a novel multistage scheme for inversion: first, we estimate the tumor initial conditions, then, given a template, we invert for scalar model parameters representing tumor proliferation, migration, and mass effect. We repeat this step for several templates and compute expectations of the observables (see §2).

2. These calculations are quite expensive. . Our solvers use efficient numerical optimization algorithms and parallelize their execution to exploit modern compute architectures. The entire method runs in parallel on GPUs so that 3D inversion on a 256^3 imaging resolution takes 1–2 hours for any subject. We open-source our parallel solver software for tumor model inversion on Github[1].

3. We use synthetic data, for which we know the ground truth, to *verify* the mathematical formulation, and convergence of our solver. This allows us to estimate the errors associated with our numerical scheme.

---

[1] https://github.com/ShashankSubramanian/GLIA

4. We apply our methods on a dataset of 216 GBM patients and evaluate our model in terms of its ability to match observed tumor patterns given the calibration: three scalar parameters and the tumor spatial initial condition. Further, we introduce a novel feature extraction methodology to obtain biophysics-based features and demonstrate their value in an important clinical task—patient stratification and overall survival prediction. We report these results in §3 and open-source our reconstructed fields in [6].

We remark that verification of the solver does not constitute validation [7] as it does not account for model errors. Verification addresses the quality of formulation, numerical schemes, and implementation, while validation focuses on the quality of the model in representing real physical processes.

B. Related work

There have been many attempts to integrate PDE models with imaging data [8]–[12]. The most common mathematical models for tumor growth dynamics are based on reaction-diffusion PDEs [13], [14], which have been coupled with mechanical models to capture mass effect [15], [16]. While there have been many studies to calibrate these models using inverse problems [11], [12], [14], [17], most do not invert for all unknown parameters (tumor initial condition and model parameters) or they assume the presence of multiple imaging scans—both these scenarios make the inverse problem more tractable. In our previous work [5], we presented a methodology to invert for tumor initial condition (IC) and cell proliferation and migration from a single scan; but we did not account for mass effect. We demonstrated the importance of using a *sparse tumor IC* (see §2) to correctly reconstruct for other tumor parameters. In [9], the authors consider a Bayesian framework for estimating model parameters but do not consider mass effect and use a single seed for tumor initial condition. The absence of mass effect in the tumor models allows for the use of simplistic precancer brain approximations through deformed templates (computed by deformable registration) [9] or simple tissue replacement strategies [18]. Other than those works, the current state of the art for single-scan biophysically-based tumor characterization is GLISTR [17].

The main shortcoming of GLISTR is that it requires manual seeding for the tumor IC and uses a single template. Further, GLISTR uses deformable registration for both anatomical variations and mass effect deformations but does not decouple them (this is extremely ill-posed), since the primary goal of GLISTR lies in image segmentation. Hence, the estimated model parameters can be unreliable due to the unknown extent to which registration accounts for mass effect and tumor shapes. Finally, the authors in [19] present a 2D synthetic study to quantify mass effect. However, they assume known tumor IC and precancer brain. To the best of our knowledge, we are not aware of any other framework that can fullyautomatically calibrate tumor growth models with mass effect for all unknown parameters in 3D. Furthermore, our solvers employ efficient parallel algorithms and GPU acceleration which enable realistic solution times, an important consideration for clinical applications.

We use a standard, single-species tumor growth model that has been used by us and other groups for algorithmic and clinical studies [9], [13], [15], [17]. This model does not capture the rich spatial inhomogeneities of observed tumors (e.g., necrosis, edema, enhancing tumor). More complex multispecies models do exist [20]–[22] but they have a large number of parameters that need calibration, which is quite challenging using single-time snapshot data. As a first step, we use a simpler model to focus on the challenges related to parameter calibration.

2. METHODOLOGY

We introduce the following notation: $c = c(\boldsymbol{x}, t)$ is the tumor concentration ($\boldsymbol{x}$ is a voxel, t is time) with *observed* tumor data $c_1 = c(\boldsymbol{x}, 1)$ and *unknown* tumor IC $c_0 = c(\boldsymbol{x}, 0)$; $m(\boldsymbol{x}, t) := (m_{\text{WM}}(\boldsymbol{x}, t), m_{\text{GM}}(\boldsymbol{x}, t), m_{\text{VT}}(\boldsymbol{x}, t), m_{\text{CSF}}(\boldsymbol{x}, t))$ is the brain segmentation into white matter (WM), gray matter (GM), ventricles (VT), and cerebrospinal fluid (CSF)[2] with *observed* preoperative segmented brain $\boldsymbol{m}_p = \boldsymbol{m}(\boldsymbol{x}, 1)$ and *unknown* precancer healthy brain $\boldsymbol{m}_0 = \boldsymbol{m}(\boldsymbol{x}, 0)$; Additionally, (κ, ρ, γ) are *calibration* scalars representing *unknown* migration rate, proliferation rate, and mass effect intensity. We model the ventricles (VT) separate from the cerebrospinal fluid (CSF) in order to inform the choice of template for the approximation of the precancerous brain anatomy (see §2-D for details).

A. Tumor growth mathematical model

Following [22], we use a non-linear reaction-advection diffusion PDE (on an Eulerian framework):

$$\partial_t c + \text{div}(c\boldsymbol{v}) - \kappa \mathscr{D} c - \rho \mathscr{R} c = 0 \text{ in } \Omega \times (0, 1] \tag{1a}$$

$$\partial_t \boldsymbol{m} + \text{div}(\boldsymbol{m} \otimes \boldsymbol{v}) = 0 \text{ in } \Omega \times (0, 1] \tag{1b}$$

$$\text{div}\left(\lambda \nabla \boldsymbol{u} + \mu\left(\nabla \boldsymbol{u} + \nabla \boldsymbol{u}^{\top}\right)\right) = \gamma \nabla c \text{ in } \Omega \times (0, 1] \tag{1c}$$

$$\partial_t \boldsymbol{u} = \boldsymbol{v} \text{ in } \Omega \times (0, 1], \tag{1d}$$

where $\mathscr{D} := \text{div } k_m \nabla c$ is a diffusion operator; $\mathscr{R} := \rho_m c(1 - c)$ is a logistic growth operator; k_m and ρ_m control the spatial heterogeneity of the diffusion and reaction coefficients in different brain tissues. Eq. (1a) is coupled to a linear elasticity equation (Eq. (1c)) with forcing $\gamma \nabla c$, which is coupled back through a convective term with velocity $\boldsymbol{v}(\boldsymbol{x}, t)$ which parameterizes the displacement $\boldsymbol{u}(\boldsymbol{x}, t)$. The linear elasticity model is parameterized by Lame coefficients $\lambda(\boldsymbol{x}, t)$ and $\mu(\boldsymbol{x}, t)$, whose values depend on the tissue type $\boldsymbol{m}(\boldsymbol{x}, t)$. We note here that $\gamma = 0$ implies $\boldsymbol{u} = \boldsymbol{0}$, which reduces the tumor growth model to a simple reaction-diffusion PDE with no mass effect (i.e, Eq. (1a) with $\boldsymbol{v} = \boldsymbol{0}$). Solution of Eq. (1) requires initial conditions for the tumor c_0 and the precancer brain anatomy $\boldsymbol{m}_0$.

---

[2]We denote CSF to be the non-ventricular cerebrospinal fluid substructure of the brain

Following [5], we parameterize $c_0 = \phi^T(\boldsymbol{x})\boldsymbol{p} = \sum_{i=1}^{m}\phi_i(\boldsymbol{x})p_i$; where $\boldsymbol{p}$ is an m-dimensional parameterization vector, $\phi_i(\boldsymbol{x}) = \phi_i(\boldsymbol{x} - \boldsymbol{x}_i, \sigma)$ is a Gaussian function centered at point $\boldsymbol{x}_i$ with standard deviation σ, and $\phi(\boldsymbol{x}) = \{\phi_i(\boldsymbol{x})\}_{i=1}^{m}$. Here, $\boldsymbol{x}_i$ are voxels that are segmented as tumor and σ is one voxel, meaning m can be quite large (~ 1000). This parameterization alleviates some of the ill-posedness associated with the inverse problem [5].

B. Inverse problem

The unknowns in our growth model are $(\boldsymbol{p}, \boldsymbol{m}_0, \kappa, \rho, \gamma)$. We calibrate $\boldsymbol{p} \in \mathbf{R}^m, \kappa, \rho, \gamma \in \mathbf{R}$ and treat $\boldsymbol{m}_0$ as a random variable on which we take expectation by solving the calibration problem for multiple $\boldsymbol{m}_0$, representing different normal templates (see numerical scheme §2-C). That is, given $\boldsymbol{m}_0$, we solve the following inverse problem:

$$\min_{(\boldsymbol{p}, \kappa, \rho, \gamma)} \frac{1}{2}\|\boldsymbol{O}c(1) - c_1\|_2^2 + \frac{1}{2}\|\boldsymbol{m}(1) - \boldsymbol{m}_p\|_2^2 + \frac{\beta}{2}\left\|\phi^T\boldsymbol{p}\right\|_2^2 \tag{2a}$$

$$\text{s.t.} \begin{cases} \mathscr{F}(\boldsymbol{p}, \kappa, \rho, \gamma) & \text{given by (1),} \\ \|\boldsymbol{p}\|_0 \leq s & \text{in } \mathbf{R}^m, \quad (2c) \\ \max(\phi^T\boldsymbol{p}) = 1 & \text{in } \Omega. \quad (2d) \end{cases} \tag{2b}$$

The objective function minimizes the L_2 mismatch between the simulated tumor $c(1) = c(\boldsymbol{x}, 1)$ at $t = 1$ and preoperative tumor data c_1 as well as the L_2 mismatch between the deformed precancer brain (due to the tumor) $\boldsymbol{m}(1) = \boldsymbol{m}(\boldsymbol{x}, 1)$ at $t = 1$ and the preoperative patient brain data $\boldsymbol{m}_p$. The mismatch terms are balanced by a regularization term on the inverted initial condition (IC). $\boldsymbol{O}$ is an observation operator that defines the clearly observable tumor margin (see §2-D for its definition; for additional details see [18]). Following [5], we introduce two additional constraints to our optimization problem—Eq. 2c parameterizes the tumor initial condition using at most s Gaussians and with Eq. 2d we assume that at $t = 0$ the tumor concentration reaches one at some voxel in the domain. Both Eq. 2c and Eq. 2d are *modeling assumptions*. We use them to deal with the several ill-posedness of the backward tumor growth PDE. In [5], we discuss this in detail.

C. Summary of our multistage inversion method

To solve Eq. 2, we propose the following iterative scheme:

(**S**.1) First, we estimate $\boldsymbol{p}$, ignoring mass effect, using the patient brain anatomy. That is, we solve Eq. 2 using the growth model with no mass effect $\mathscr{F}(\boldsymbol{p}, \kappa, \rho, \gamma = 0)$ for $(\boldsymbol{p}, \kappa, \rho)$.[3] Our *precancer* scan is approximated by $\boldsymbol{m}_0 = \boldsymbol{m}_p$ with tumor regions replaced by white matter.

(**S**.2) Next, we approximate $\boldsymbol{m}_0$ using a normal template scan $\boldsymbol{m}_t$. Since $\boldsymbol{p}$ is defined in $\boldsymbol{m}_p$, we register $\boldsymbol{m}_p$ to $\boldsymbol{m}_t$ and transfer $\boldsymbol{p}$ to the $\boldsymbol{m}_t$ space[4]. This registration is used to

---

[3]$\gamma = 0$ indicates no mass effect and Eq. (1b), Eq. (1c) and Eq. (1d) are not needed

avoid, to the best extent possible, the tumor initial locations in anatomical structures such as ventricles (where the tumor is not known to grow).

(**S**.3) Finally, we solve Eq. 2 for κ, ρ, and γ using the estimated p and approximation $m_0 = m_t$.

We repeat (**S**.2) and (**S**.3) for an ensemble of templates m_t to make our inversion scheme less sensitive to the template selected. In (**S**.1), we employ the fast *adjoint-based* algorithm outlined in [5] that alternates between an ℓ_0 projection to enforce sparsity and an ℓ_2 regularized solve for the unknowns to obtain a *sparse* and *localized* tumor initial condition. For the ℓ_2 regularized solve in (**S**.1) and the inverse solve in (**S**.3), we use a quasi-Newton optimization method (LBFGS) globalized by Armijo linesearch with gradient-based convergence criteria. For (**S**.3) we employ first order finite differences to approximate the gradient since there are only three scalar unknowns.

D. Selection of template and numerical parameters

We observe variability in estimated parameters across different m_t. We were able to reduce this variability by selecting m_t that are more similar to m_p using ventricle size as similarity. Since tumor mass effect typically compresses the ventricles, we select those templates that have larger ventricular volumes than the patient and within a relative margin of 30%. This margin (30%) is computed by simulating several large tumors with strong mass effect and close proximity to the ventricles to estimate the extent of mass-effect-induced ventricle compression. Here, we assume that the ventricles are a *qualitative indicator* of patient-template anatomical similarity. More accurate similarities are problematic to compute as they rely on image registration which is unable to decouple tumor mass effect from anatomical differences.

Next, we discuss the selection of additional parameters used in our solver. These parameters are default and do not require any subject-specific tuning.

    i.   We use 80 normal subject scans from the Alzheimers Disease Neuroimaging Initiative (ADNI)[5] normal/control subject dataset as our sample space for templates. In order to get good expectation estimates, it is preferable to use a large number of templates. We choose 16 templates for two reasons: First, we were unable to consistently[6] find a larger number of templates that satisfy our preselection criteria across the 216 GBM patient clinical dataset, and second, the sensitivity of inversion to the number of templates does not alter (drop) significantly with larger number of templates (see discussion on sensitivity to number of templates in appendix Tab. A2).

    ii.   Experimental evidence suggests that glioma cells infiltrate faster in white matter [23]. We choose $\kappa_{\text{GM}} = 0.2\kappa_{\text{WM}}$ where κ_{GM} and κ_{WM} are the tumor diffusion rates in the gray matter and white matter, respectively. Literature [9], [13], [17] suggests

---

[4]we simply mask the tumor region for this registration
[5]ADNI database adni.loni.usc.edu
[6]greater than 50% of the patient cohort is able to yield sufficient number of candidate templates based on our selection criterion

values of 0.1 and 0.2. We do not observe any significant sensitivity of our inversion to either choice (also observed in the independent study in GLISTR [17]).

iii. Elastic properties for different tissue types are fixed based on [22]. See appendix Tab. A1.

iv. Optimizer convergence parameters (tolerance, maximum iterations) are selected based on [18] using tests which balance run times and solution accuracy on synthetic datasets (see §3 and appendix Tab. A1).

v. The standard deviation σ of the Gaussian functions is taken as the resolution of our input MRI image. We select Gaussians from an equally-spaced[7] candidate set of Gaussians at voxels that are segmented as tumor. These form our basis functions $\phi_i(\boldsymbol{x})$ (see §2-A) for the tumor IC c_0.

vi. The sparsity level s of the tumor IC is based on an empirical study on a large clinical cohort [18]—the study observed values between 3 and 6 to show a good compromise between a localized sparse solution and the reconstruction of complex tumor shapes (as measured by the dice coefficient). We select s to be 5. Note that s is an upper bound on expected number of activation sites for the tumor IC; the actual number of sites is determined by solving Eq. 2.

vii. We define our observation operator application as $\boldsymbol{O}c(\boldsymbol{x}, 1) := o(\boldsymbol{x}) \odot c(\boldsymbol{x}, 1)$ (see (2)). Here, $\odot$ signifies hadamard product and $o(\boldsymbol{x})$ is defined to be one in every voxel $\boldsymbol{x}$ except the peritumoral edema (ED), where it is zero. That is, we observe the tumor concentration c everywhere except within the edema where it is set (observed) as zero. We make this choice because the peritumoral region in GBMs is known to be infiltrated with cancer cells to an *unknown* extent [24]. The observation operator can be interpreted as complete uncertainty of tumor observation in this region. Hence, the tumor profile in the peritumoral edema is governed solely by the model dynamics, and not fitted to arbitrary values (see [18] for more details).

E. Biophysics-based feature extraction

Here we introduce a set of manually crafted features that are solely based on the output of the biophysical model and can potentially be used in downstream clinical tasks. These features are highly correlated since they are generated by our model and are, hence, based on three scalar biophysics parameters, the tumor initial condition, and the underlying brain anatomy. We pose the question as to whether such features can be useful and evaluate this on an important clinical task, namely patient survival prediction, in §3-B.

1.   Spatial extraction: We extract features related to tumor concentration, displacement, and stresses at various times of the tumor evolution, including *future* times. We compute the

---

[7]The spacing between Gaussians is $\delta = 2\sigma$, which is empirically obtained to guarantee good conditioning of $\int_{\Omega} \phi\phi^T dx$

following spatial fields: c, $\partial c/\partial t$, $\| \nabla c \|_2$, $\| u \|_2$, trace(T), τ, $\mathscr{I}: = \det(I + \nabla u)$, $\| v \|_2$, where $T: = \lambda \text{div } uI + \mu(\nabla u + \nabla u^T)$ is the elasticity stress tensor, τ is the maximum shear stress field[8], det is the determinant, and $\mathscr{I}$ is the deformation Jacobian. Then, given any of these scalar fields $y(x)$, we compute the following features: $\int_\Omega y d\Omega$, standard deviation (quantifies spread), 87.5^{th}, 95th, and 99th percentiles (quantifies magnitudes), and surface integral $\int_\Gamma y d\Gamma$ at the isosurface $c = c^\star$ for prespecified $c^\star$, which we approximate by:

$$\int_\Gamma y(x)d\Gamma = \frac{N^{1/3}}{2\pi} \int_\Omega \mathbf{1}\big(\big|c(x) - c^\star\big| < \epsilon\big) \odot y(x)dx \tag{3}$$

for some tolerance ϵ[9]. Here, N is total number of voxels, $\mathbf{1}$ is an indicator function, and $\odot$ indicates hadamard product. To compute $c^\star$, we first calculate the segmentation of the brain—each voxel is assigned the label (healthy tissues or tumor) with highest concentration value. Then, we compute $c^\star = c(x^\star)$, where $x^\star = \text{argmax}_{x \notin \text{tumor}}(\| \nabla c \|_2)$; this is the isosurface at the tumor front where the tumor shows strong gradients and represents the evolving front. The volumetric features are computed for different spatial regions: healthy brain $\mathscr{B}$ (that is, every non-tumorous brain voxel), the tumor region, and the full spatial domain Ω. In total, we compute 128 features at each time t.

2. Temporal extraction: While the exact time point is impossible to ascertain, we compute these features at five different *pseudotime* points: these are the diagnosis/imaging scan time and the time instances at which volume fraction of the tumor is 0.1%, 1%, 3%, and 5% of the brain. We have chosen these times to *temporally standardize* all our features, since every patient (at diagnosis) might be at a different stage of cancer evolution. For example, for patients with large tumors, the features are mostly computed at past times, and for patients with smaller tumors, they represent forecasts (assuming no treatment). The biophysical model makes such forecasts possible. .

F. Workflow

Here we summarize the workflow we used to process a clinical subject dataset.

1) Image preprocessing: Our dataset consists of single-time-point (preoperative) structural mpMRI scans (T1, T1-CE, FLAIR, T2) of 216 GBM patients, along with age, gender, surgical resection status, and overall survival information. For each patient, the four mpMRI scans are first affinely registered to a template atlas. Then, they are segmented into tumor regions of enhancing tumor (ET), necrotic tumor (NEC), and peritumoral edema (ED) using a state-of-the-art neural network [25] trained on the BraTS [26] dataset which provides expert radiologist ground truth labels for these tumor regions. Since our tumor model is a single-species growth model, we define the tumor region/label as the tumor core (TC) := ET∪NEC. Finally, they are further segmented into healthy substructures (GM,

---

[8] τ can be computed as $1/2(e_3 - e_1)$, with e_1 and e_3 as the minimum and maximum eigenvalues of the stress tensor T

[9] ϵ is computed experimentally as 0.1 to ensure a continuous but sufficiently narrow surface

WM, VT, and CSF) using ANTs [27], where we register several templates to any given patient and ensemble the atlas-based segmentation for the different healthy tissue labels. The segmentation labels serve as a proxy for our tissue (and tumor) concentration data.

2) Inversion software: Our open-sourced software is written in C++ and parallelized using MPI and CUDA. The solver is capable of running on several CPUs as well as GPUs (a single GPU suffices for medical imaging resolutions). A forward solve of our mass effect model on 256^3 spatial dimensions takes less than 2 minutes on a V100 NVIDIA GPU and the full inversion takes an average of 1–2 hours (depending on the input tumor data and number of m_t templates in the ensembled inversion). We refer the reader to [5], [28] for our solver numerics and parallel algorithms.

3. RESULTS

A. Synthetic verification

In [4], we verified the correctness of the solver using synthetic data. We briefly summarize the tests and results here. We ask the following three questions:

(**Q1**) Given p (tumor IC) and m_0 (precancer scan), can we reconstruct (κ, ρ, γ) (i.e., the diffusion, reaction, and mass effect coefficients) using (**S**.3) from §2-C?

(**Q2**) Given m_0 but unknown p, can we reconstruct $(p, \kappa, \rho, \gamma)$ using (**S**.1) and (**S**.3)?

(**Q3**) With m_0 and p unknown, can we reconstruct $(p, \kappa, \rho, \gamma)$ using (**S**.1)-(**S**.3), using ensembled inversion with $m_0 = m_t$?

To create the synthetic data, we select m_0, p, κ, ρ, and γ and run our forward solver to $t = 1$. The model outputs at $t = 1$ define our synthetic data which is then used to reconstruct $(p, \kappa, \rho, \gamma)$, after which we compare the errors. We consider the following test-cases (variations) for our parameter configurations:

$$TC(a): \quad \text{no mass effect} \quad \left(\gamma^\star, \rho^\star, \kappa^\star\right) = (0, 12, 0.025)$$

$$TC(b): \quad \text{mild mass effect} \quad \left(\gamma^\star, \rho^\star, \kappa^\star\right) = (0.4, 12, 0.025)$$

$$TC(c): \quad \text{medium mass effect} \left(\gamma^\star, \rho^\star, \kappa^\star\right) = (0.8, 10, 0.05)$$

$$TC(d): \quad \text{large mass effect} \quad \left(\gamma^\star, \rho^\star, \kappa^\star\right) = (1.2, 10, 0.025)$$

where $\star$ represents the non-dimensionalized *ground truth* parameters. The tumors along with the deformed m_0 anatomy are visualized in Fig. 1.

We report relative errors for κ, ρ, and γ. We compute displacement two-norm field $u = u(\boldsymbol{x}) = \parallel \boldsymbol{u}(\boldsymbol{x}) \parallel_2$ (this field quantifies mass effect deformation) and report the relative error in the norm $\parallel u \parallel_2$ and the infinity norm $u_\infty = \parallel u \parallel_\infty$ (in mm). Finally, we report the dice coefficient in reconstructing the tumor core π_{TC}.

(**Q1**) *Known $\boldsymbol{m}_0$, known $\boldsymbol{p}$*: We report the reconstruction errors in Tab. I under "True IC". We observe excellent reconstruction for all test-cases with relative errors less than 2%.

(**Q2**) *Known $\boldsymbol{m}_0$, unknown $\boldsymbol{p}$*: We use our inversion scheme outlined in §2-C, where we first invert for the tumor initial condition using (**S**.1) and then the model parameters with *known* precancer scan. We present the reconstruction errors in Tab. I under "Inverted IC". The errors increase because we reconstructed $\boldsymbol{p}$ using the no-mass-effect model and the inverse problem is severly ill-posed (see [5] for analysis of the non-convexity). But we can still recover the model parameters quite well: the mass effect (error in two-norm of the displacement norm e_u) is captured with relative errors less that 8%; the reaction and diffusion coefficient also have good estimates of around 16% and 17% average relative errors, respectively. We show the estimated tumor initial conditions for a specific test-case in Fig. 2 using (**S**.1). We observe that the reconstructed initial conditions are in a localized neighbourhood around the ground truth initial condition for all test-cases. Due to the inherent ill-posedness and model approximation errors, the exact initial locations are not recoverable.

(**Q3**) *Unknown $\boldsymbol{m}_0$, unknown $\boldsymbol{p}$*: This scenario corresponds to the actual clinical problem. We invoke (**S**.2) and use an ensemble of 16 templates to approximate $\boldsymbol{m}_0$ (see §2-D). We report inversion results in Tab. II and show an exemplary reconstruction of the patient using the different templates in Fig. 3. We report the expectations (mean; μ_ι) and standard deviations (σ_ι; characterizes the sensitivity of our inversion to each template) for each parameter $\iota = (\gamma, \rho, \kappa, u_\infty)$. We refer to the tumor-deformed template ($\boldsymbol{m}(\boldsymbol{x}, 1)$) with the grown tumor ($c(\boldsymbol{x}, 1)$) using the reconstructed parameters as the *reconstructed patient*.

For the tumor dice coefficient, we compute the probability $p_l(\boldsymbol{x})$ for any tissue/tumor label l at location $\boldsymbol{x}$ as:

$$p_l(\boldsymbol{x}) = \frac{\sum_{t=1}^M \mathbf{1}(s_t(\boldsymbol{x}) = l)}{M}, \tag{4}$$

where $s_t(\boldsymbol{x})$ is the label map, M is the total number of templates (16 in our case), and $\mathbf{1}$ is an indicator function. Then, we construct a *probabilistic label image* by simply assigning every voxel to the label with maximum probability. The tumor dice is computed using the tumor label from the probabilistic reconstructed patient and the patient tumor data.

Despite not knowing the true $\boldsymbol{m}_0$, we are still able to capture the parameters. The expectation of the displacement norm error has an average error of about 17%, while the expectations of reaction and diffusion coefficients show about 24% and 15% average relative errors, respectively. The reaction coefficient shows least sensitivity (standard deviations are an

order of magnitude smaller than the expectations); the diffusion coefficient also shows small sensitivity which is roughly the order of numerical diffusion that is introduced in our forward solvers[10]; the maximum displacement shows standard deviations of the order of 1 mm. The mass effect parameter γ shows high standard deviations for the test-cases with small mass effect and progressively improves as the mass effect increases. This is due to ill-conditioning with respect to γ when either the tumor or the mass effect are small.

B. Clinical data analysis

We apply our inversion scheme on the 216 clinical GBM patient cohort (see §2-F for details regarding the dataset and preprocessing steps). We show some example initial condition reconstructions for a few GBM subjects (with multifocal examples) in Fig. 4. The solver reconstructs sparse and localized tumor initial conditions (irrespective of the focality of the tumor data) across the clinical cohort. While certain initiation locations are close to the center of mass, we observe additional initiation sites are required to capture the complex hetereogeneous shape of brain tumors. We report exemplary reconstruction results from the inversion scheme in Tab. III for four patients with different ranges of mass effect. We observe similar trends in sensitivity (standard deviations) of parameters and dice coefficients as in the synthetic test-cases.

We visualize the reconstructions in Fig. 5. We show the averaged[11] probabilistic template image, the averaged reconstructed patient image, the patient data, and the average reconstructed displacement two-norm field. We observe good quantitative agreement of the tumor (dice coefficients greater than 83%) and good quantitative characterization of the mass effect. Patients *AAUJ* and *AANN* present large tumors with significant mass effect ($\| u \|_\infty \sim$ 18mm $\pm$ 1 and 12mm $\pm$ 5, respectively); patient *AARO* also shows a large tumor but with mild mass effect ($\| u \|_\infty \sim$ 7mm $\pm$ 1); patient *AALU* presents a small tumor with negligible mass effect($\| u \|_\infty \sim$ 2mm $\pm$ 1). In Fig. 6, we show results stratified according to the extent of mass effect.

In order to understand the role and importance of mass effect, we conduct a preliminary analysis on the patient cohort using two different growth models: model *ME* represents a growth model with mass effect and model *NME* represents a model without mass effect (i.e., $\gamma = 0$ or a simple nonlinear reaction-diffusion PDE). We compare the two using the three questions below:

(**CQ1**) Does model *ME* improve the reconstruction dice coefficients?

(**CQ2**) Is mass effect (as quantified through model *ME*) simply a feature of larger tumors?

(**CQ3**) Do features extracted from our biophysical model (see §2-E) add value to important clinical evaluations?

---

[10]We employ spectral solvers that require some artificial diffusion through smoothing to prevent aliasing errors; see appendix for more details
[11]Averaging results in coarse representations of the ensemble and is, hence, only used for visualization purpose

(**CQ1**) We plot the dice coefficients (see (4)) for the tumor core (TC) and ventricles (VT) in Fig 7a. We sort the results in increasing order of mass effect (using the reconstructed u_∞ and $\| u \|_2$); the hypothesis is that **ME** is more informative as mass effect becomes a pronounced feature of the disease. We observe that the tumor dice coefficients (π_{TC}) show similar performance for both **ME** and **NME**. The average TC dice coefficients across the population are 0.849 ± 0.063 and 0.857 ± 0.073 for **ME** and **NME**, respectively. This is not surprising since **NME** selects p, κ, and ρ to match the observed tumor core and does not consider matching the anatomy. In contrast, **ME** has to account for both the anatomy and the tumor. For the VT dice, we observe that **ME** shows an overall better performance that progressively improves as the mass effect increases. Across the population, the average VT dice coefficients are 0.621 ± 0.091 and 0.563 ± 0.120 for **ME** and **NME,** respectively—an approximate 6% performance improvement. For patients whose displacement norms are greater than the median displacement norms (median is 6.76 mm; we consider these patients to qualitatively have moderate to large mass effect), this improvement is approximately 10%.

Here, we note that the VT dice coefficients are low due to the significant anatomical variability between the templates and the patient. Further, many subjects may show other neurodegenerative disease patterns like abnormal ventricles due to Alzheimer's disease that our model does not take into account, leading to larger template-patient anatomical variations. However, this can be remedied with a subsequent image registration between the reconstructed patient and the patient data—the deformation from mass effect has been *decoupled* through our inversion scheme and, now, registration can be used to correct for anatomical differences.

We also plot the variability of the biophysical scalar parameters in Fig 8. We observe that both models provide a *different* set of reaction and diffusion coefficients. This is because **NME** captures the tumor but ignores mass effect, which leads to larger κ and ρ. In contrast, **ME** accounts for both growth and anatomy deformation. Interestingly, the fitted parameters exhibit smaller cross-subject variability compared to **NME**.

In Fig 9a, we compare the two models. Subject *AAVD* is a multicentric GBM with negligible mass effect; both models show similar dice scores ((TC, VT) dice scores of (0.795, 0.613) and (0.813, 0.641) for **ME** and **NME**, respectively); subject *AAUX* has significant mass effect (as seen by the compressed ventricles and mid-line shift) and **ME** outperforms **NME** in its reconstructions ((TC, VT) dice scores of (0.84, 0.527) and (0.842, 0.329) for **ME** and **NME**, respectively), with nearly 20% improvement in VT dice coefficients and similar TC dice coefficients. We remark that this improvement is achieved by only one additional scalar parameter (γ) in our biophysical model.

Finally, in Fig 10 and Fig. A.2 and Fig. A.3, we expand on the dice coeffcient values and plot the variability of the dice coefficients using the **ME** model to understand the distribution and source of outliers. We observe some outliers in our reconstructions where the tumor (TC) dice coefficient π_{TC} is less than 0.70. We find that these outliers can be linked to specific reasons (see Fig. A.2 and Fig. A.3 for specific examples): first, the template and patient anatomical difference can be significant (such as very dissimilar ventricle structures) leading to low quality tumor reconstructions, second, the tumor segmentation itself may not

be accurate, and third, mass effect induced by edema is not accounted for in our models and this could lead to unrealistically large tumor core volumes to match the observed deformations.

(**CQ2**) In Fig. 7b we report the patient tumor (TC) and edema (ED) volumes; point color and size correspond to the mean displacement field μ_u. We observe that tumors with similar volume can exhibit a wide range of mass effect. For instance, tumors with volume around 70–80 cm^3 show deformations in the negligible range as well as substantial mass effect (> 15 mm). In Fig 9b, we show tumors with volumes within 10% of each other, but mass effect displacements ranging from 4 mm to 17 mm (also qualitatively visible in the reconstructed displacement two-norm fields). Thus, we conclude that mass effect is not a feature of larger tumors and can possibly be an additional and valuable biomarker for GBMs.

(**CQ3**) We consider survival data for patients that underwent a gross total surgical resection (GTR)—our survival dataset consists of 141 patients (GTR only) and each patient is classified as a short, mid, or long survivor based on equal quantiles of survival time (in days). Our objective is to use patient-specific features to predict the survival class by training a machine learning classifier on our dataset. We consider two groups of features: *age and volumetric features* that include patient age and tumor volumes for each tumor subclass, and *biophysics-based features* that are extracted using our calibrated growth model (see §2-E; we show an example visualization of a few biophysics-based features in Fig 11).

    i.   *Feature correlations:* We plot the Spearman correlation of statistically significant feature correlations with survival time in Fig. 12. We observe that age is the maximally correlated feature with survival—an observation consistent with several previous studies [29], [30]. We also note that biophysics-based features correlate better than volumetric features (tumor volumes; correlations are insignificant and small), and features computed at standardized times (at specific tumor volume fractions) correlate better than features computed at diagnosis/ imaging scan time.

    ii.   *Univariate patient stratification:* We stratify patients based on the median values of a few top correlated features (with survival) and plot the Kaplan-Meier (KM) survival curves for each feature (see Fig. 13). We observe that several biophysics-based features (along with age) dichotomize the patients based on their overall survival. For instance, patients with large values of trace(T) (relates to the tumor-induced pressure due to mass effect) show reduced overall survival. We also additionally show the stratification based on a volumetric feature (tumor core volume). We observe volumetric features are unable to stratify the subject survival curves.

    iii.   *Survival classification:* We employ random forests as our survival classification model on an 80–20 train-test data split. We perform a univariate feature selection where we first sort features according to their average 5-fold cross-validated (CV) random forest feature importance. Then, we compute a cross-validated (CV) training accuracy and standard deviation (std) by adding features incrementally in descending order of importances (see Fig. 14). We then select

the top 13 features (these include age, edema volume, and other biophysics-based features) since they show maximum CV accuracy and small accuracy standard deviations across the splits (hence prone to lesser overfitting). We report our classification scores in Tab. IV. We observe an improvement in classification scores using the selected features in comparison to using only age and volumetric features. The classifier also shows smaller variability across cross-validated splits and hence, potentially, generalizes well (as suggested in the testing accuracies).

To bolster our argument, we also repeat our inversion and survival classification on the publicly available BraTS 2020 challenge dataset [26]. We submit our classification results to the BraTS 2020 leaderboard validation portal and report the accuracies in Tab. V. We note the following results: First, we observe similar trends of accuracy improvements using biophysics-based features on an 8020 split of the BraTS training dataset—the biophysics-based features show improved performance over age and volumetric features on cross-validated training and testing accuracies. Second, these features also show improved accuracies on the challenge leaderboard (unseen) validation dataset of about 6% compared to age and volumetric features. Finally, the leaderboard accuracies are comparable to the top teams of the BraTS challenge[12], demonstrating that the proposed methodology is competitive.

4. CONCLUSIONS

Our results are very promising. First, our solver is robust (never crashes, takes excessive iterations, or needs subject-specific hyperparameter settings). It does not require any manual preprocessing and can be run in a black box fashion. Our experiments on synthetic data with known ground truth is the first time that such a solver is verified and we demonstrated that our approximations do not introduce significant errors despite the fact that m_0 is unknown. For clinical data, the model errors are expected to dominate the errors from using a template and splitting the calibration procedure into two stages. Second, we tested our method on a large clinical patient cohort (216 patients) in order to test the feasibility of our method. With a small number of calibration parameters, our solver is able to quantitatively match the observed tumor margins and qualitatively correlate with observed mass effect. Hence, our solver provides a means to quantify and localize mass effect without relying on any assumptions on symmetry and location of the tumor. Further, we showed that including mechanical effects in our tumor growth models can provide additional information in terms of capturing the deformations of the tumor and surrounding anatomical structures and results in quantitatively different parameters and reconstructions than a non-biomechanical growth model. Finally, we introduced a novel biophysics-based feature extraction method and demonstrated the clinical value of the extracted features on an important clinical task—patient stratification and survival prediction using pretreatment imaging scans. While age remains an important population-based feature for predicting survival outcome, our observations indicate that biophysics-based features can enhance the quantification of survival prognosis. Specifically, these features closely follow the predictive capability of patient age in the following: first, correlation with patient overall

---

[12]The first rank team's validation scores were not reported. Hence, we compare with the second and third rank teams

survival, second, dichotomization (stratification) of patients based on overall survival, and last, improving prediction of patient overall survival through machine learning methods. Biophysics-based features are also competitive with the winning algorithms of the BraTS challenge competition. However, we do note the low survival classification accuracies, indicative of the difficulty in using machine learning methods on small datasets (141 GTR patients here). Similar observations have been made before on the BraTS dataset for survival prediction [26].

Some of the *limitations* of our work are as follows. Our forward model uses a single tumor species that cannot distinguish between the enhancing, necrotic, and tumor-infiltrated edema regions of the brain. While our survival analysis is encouraging, it is preliminary—other factors such as genetic information (IDH1/EGFR mutation, MGMT promoter methylation), radiotherapy dosage profiles, postoperative recurrence information are important prognostic indicators of survival. A comprehensive clinical study would need to include such data and is the focus of our future work. Further, in this study, we are limited to structural mpMRI scans. Other imaging sources such as Diffusion Tensor Imaging scans or Perfusion scans could potentially be useful. Our next step is to incorporate multiple tumor species in our growth model and continue our clinical evaluation using a larger patient cohort with additional genetic information. Our hope is that such models can also assist in additional clinical tasks such as tumor recurrence prediction.

A.: APPENDIX

Model and inversion additional parameter values:

We report additional solver parameter values (based on [22]) in Tab. A.1. There is no subject-specific tuning of any parameter. The tumor initial condition (IC) reconstruction (**S**.1) is performed using a spatial resolution of 256^3, to ensure smallest possible IC. The ensemble inversion for parameters (**S**.3) is performed in a downsampled resolution 160^3 for faster solution times. We do not observe significant parameter inversion sensitivity between the two resolutions for the final step. If (**S**.3) is also performed in 256^3, the full inversion time is about 4–6 hours on a V100 GPU. Our current inversion time is about 1–2 hours. Finally, for the ensemble inversion, we observe that in rare cases some templates produce potentially unreliable solutions—we discard solutions with large condition number (> 1000) of the finite difference 3×3 Hessian matrix (there are three scalar unknowns—κ, ρ, γ) of the objective function.

Forward solver numerics:

We use a pseudo-spectral Fourier method on a regular mesh for spatial discretization and employ 3D fast Fourier transforms to compute all spatial differential operators. We enforce our boundary conditions on the surface of the brain and VT/CSF boundaries using a penalty approach (see [11], [12]). We employ no-flux boundary conditions on our state variables at the brain and VT/CSF boundaries. Further, we use zero displacement boundary conditions at the brain boundary. We use periodic boundary conditions on $\Omega = [0, 2\pi]^3$. We use the Strang operator splitting scheme to solve the reaction, advection, and diffusion splits. The reaction

split is solved analytically, the diffusion split is solved using the Crank-Nicolson scheme through a preconditioned Conjugate Gradient method, and the advection split is solved using the semi-Lagrangian scheme. The variable coefficient elasticity equations are also solved using the preconditioned Conjugate Gradient method. We refer the reader to [22] for further details on the numerical schemes and convergence.

TABLE A.1:

Model and inversion additional parameter values used in our simulations. Note that the Lamè coefficients λ and μ are determined by the Young's modulus and Poisson's ratio of the tissue-type.

Parameter	value
Young's modulus of (GM, WM, CSF, tumor) (Pa)	(2100, 2100, 100, 8000)
Poisson's ratio of (GM, WM, CSF, tumor)	(0.4, 0.4, 0.1, 0.45)
Gaussian width, σ (voxels)	1
Spacing between Gaussians δ (voxels)	2
Regularization parameter, β	1
Relative gradient tolerance	1E-3
Relative objective function change tolerance	1E-3
Maximum number of quasi-Newton iterations	50

Inversion additional results:

We provide some additional reconstruction details and experiments:

i.   We report our inversion results using two choices for number of templates (8 and 16) for precancer brain m_0 approximation in Tab. A.2. We observe insignificant sensitivity—the standard deviation of our reconstructed parameters and metrics are representative of the anatomical variation (and approximation error) between the patient and the chosen template.

ii.   We observe that the inclusion of ventricles (VT) in the objective function (in the data misfit term $\| m(1) - m_p \|_2^2$ is important. While it does not make a significant difference for several patients, it does lead to different (and qualitatively inaccurate) results for certain patients. We compare our reconstructions in Fig. A.1 and Fig. A.3 for one such patient. We observe that the absence of the VT misfit term leads to erroneous reconstruction of mass effect (the average displacement norm is almost halved). This can be observed in our visualization as well as the reconstructed parameters and dice coefficient for VT. Another note: we do not include GM, WM, and CSF mismatches in our objective function (only tumor and VT included) since their segmentations (from ANTs) can be poor and they also show high anatomical dissimilarity.

iii. We visualize reconstructions for outlier cases (low tumor dice scores) in Fig. A.2: we observe that high anatomical variations between the template and the patient can contribute significantly to low dice scores. Further, in some cases (such as patient AAXV), the mass effect may be (qualitatively) due to edema which our model does not capture. Hence, the inversion results in larger tumor core volumes to account for the deformation. Finally, in Fig. A.3, we show an alluvial plot that higlights the correlations between tumors with small mass effect and outlier dice scores (in tumor core and ventricles).

Machine learning classifier details:

We employ a random forest classifier for our survival prediction task. The hyper-parameters of the random forest include maximum tree depth (taken as 3) and number of estimators (taken as 100), and are estimated during our cross-validation procedure. Finally, we additionally experimented with other classifiers and regressors (logistic regression, linear regression, and elastic nets) but observed that • random forests provide the highest classification accuracy and • similar conclusions from the main text can be drawn from a regression analysis.

TABLE A.2:

Inversion results sensitivity with number of templates for clinical data. All results are repeated for two choices of number of templates—8 and 16. We report the mean (μ_ι) and standard deviation (σ_ι) of the reconstructed parameters across the ensemble of templates; $\iota = \gamma, \rho, \kappa$, and the maximum value of the displacement field two-norm (i.e., $u_\infty = \| u \|_\infty$ in mm). We also report the probabilistic tumor core dice coefficient π_{TC}, the average runtime μ_T, and standard deviation in runtime σ_T in hh:mm:ss. (Note: the diffusion coefficient (κ) values are scaled by $1E-2$ for clarity; Patient AANN resulted in less than 8 templates for preselection due to the abnormally large ventricles—the inversion is identical for both template number choices).

| Patient ID	Num. of templates	$(\mu_\gamma, \sigma_\gamma)$	(μ_ρ, σ_ρ)	$(\mu_\kappa, \sigma_\kappa)/0.01$	$(\mu_{u_\infty}, \sigma_{u_\infty})$	π_{TC}	(μ_T, σ_T) (in hh:mm:ss)
AAUJ	8	(0.81, 0.065)	(9.83, 0.83)	(2.09, 1.14)	(17.17, 1.05)	0.871	(01:13:51,00:18:27)
AAUJ	16	(0.82, 0.081)	(9.92, 0.75)	(1.87, 0.89)	(17.57, 1.4)	0.875	(01:13:27,00:13:43)
AANN	8	(0.59, 0.25)	(9.93, 0.61)	(2.28, 1.23)	(11.68, 4.91)	0.853	(01:27:24,00:06:23)
AANN	16	(0.59, 0.25)	(9.93, 0.61)	(2.28, 1.23)	(11.68, 4.91)	0.853	(01:27:24,00:06:23)
AARO	8	(0.42, 0.109)	(10.54, 0.28)	(1.8, 0.37)	(7.42, 1.79)	0.833	(01:08:03,00:07:09)
AARO	16	(0.4, 0.089)	(10.64, 0.58)	(1.75, 0.38)	(7.23, 1.49)	0.839	(01:08:19,00:05:47)
AALU	8	(0.51, 0.22)	(6.34, 0.45)	(0.97, 0.55)	(3.36, 1.36)	0.928	(01:06:52,00:07:55)

Patient ID	Num. of templates	$(\mu_\gamma, \sigma_\gamma)$	(μ_ρ, σ_ρ)	$(\mu_\kappa, \sigma_\kappa)/0.01$	$(\mu_{u_\infty}, \sigma_{u_\infty})$	π_{TC}	(μ_T, σ_T) (in hh:mm:ss)
AALU	16	(0.36, 0.23)	(6.39, 0.52)	(1.1, 0.55)	(2.43, 1.48)	0.942	(01:05:22,00:08:14)

TABLE A.3:

Inversion results for patient AAUJ by including ventricle data misfit (with VT) and excluding (no VT) it in the objective function. We report the mean (μ_ι) and standard deviation (σ_ι) of the reconstructed parameters across the ensemble of templates; $\iota = \gamma, \rho, \kappa$, and the maximum value of the displacement field two-norm (i.e., $u_\infty = \| u \|_\infty$ in mm). We also report the probabilistic tumor core dice coefficient π_{TC}, the probabilistic ventricle dice coefficient π_{VT}, the average runtime μ_T, and standard deviation in runtime σ_T in hh:mm:ss.

Patient ID	Objective function	$(\mu_\gamma, \sigma_\gamma)$	(μ_ρ, σ_ρ)	$(\mu_\kappa, \sigma_\kappa)/0.01$	$(\mu_{u_\infty}, \sigma_{u_\infty})$	π_{TC}	π_{VT}	(μ_T, σ_T) (in hh:mm:ss)
AAUJ	no VT	(0.38, 0.19)	(10.99, 0.75)	(1.73,0.76)	(8.38, 4.08)	0.926	0.461	(01:08:45,00:06:04)
AAUJ	with VT	(0.82, 0.081)	(9.92, 0.75)	(1.87,0.89)	(17.57, 1.4)	0.875	0.598	(01:13:27,00:13:43)
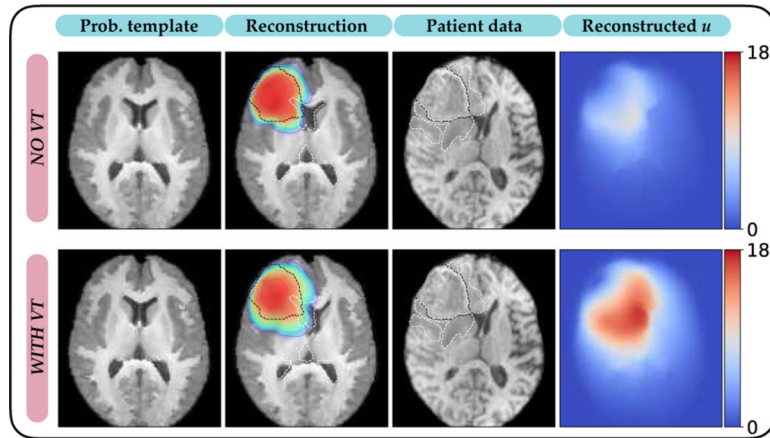
Fig. A.1:
Reconstruction using two types of objective functions—with VT (VT misfit included) and no VT (only tumor misfit). The first column shows the averaged template T1 MRI, the second column shows the averaged patient reconstruction with the estimated tumor concentration (higher concentrations (~1) are indicated in red and lower ones in green) overlaid. We also show the underformed ventricle contour (white dashed line) to express the extent of mass effect, as well as the TC data contour (black dashed line). The third column shows the patient T1 MRI data with the TC data contour (black dashed line) and the edema data contour (white dashed line). The last column depicts the reconstructed average displacement (in mm) two-norm field. We observe different reconstructions—the no VT result shows much smaller mass effect and deformation of ventricles and is qualitatively inaccurate.
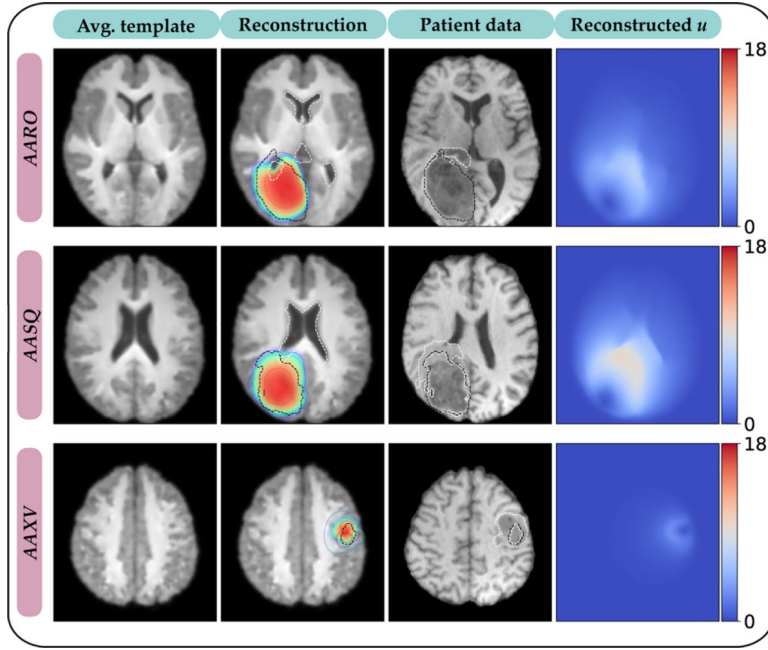
Fig. A.2:

Reconstructed images using **ME** model for three patients with outlier tumor (TC) dice coefficient (π_{TC}). The first column shows the averaged template T1 MRI, the second column shows the averaged patient reconstruction with the estimated tumor concentration (higher concentration (~1) are indicated in red and lower ones in green) overlayed. We show underformed ventricles contour (white dashed line) to express the extent of mass effect as well as the TC data contour (black dashed lines). The third column shows the patient T1 MRI data with the TC data contour (white dashed line). The last column depicts the reconstructed average displacement (in mm) two-norm field. We observe π_{TC} values as 0.650, 0.657 and 0.667 for patients AARO, AASQ and AAXV, respectively. We observe high anatomical difference in ventricles between the average template and patient data for patient AARO, qualitative inaccurate tumor segmentation for patient AASQ (i.e. detecting a section of edema as TC) and mass effect due to edema for patient AAXV.
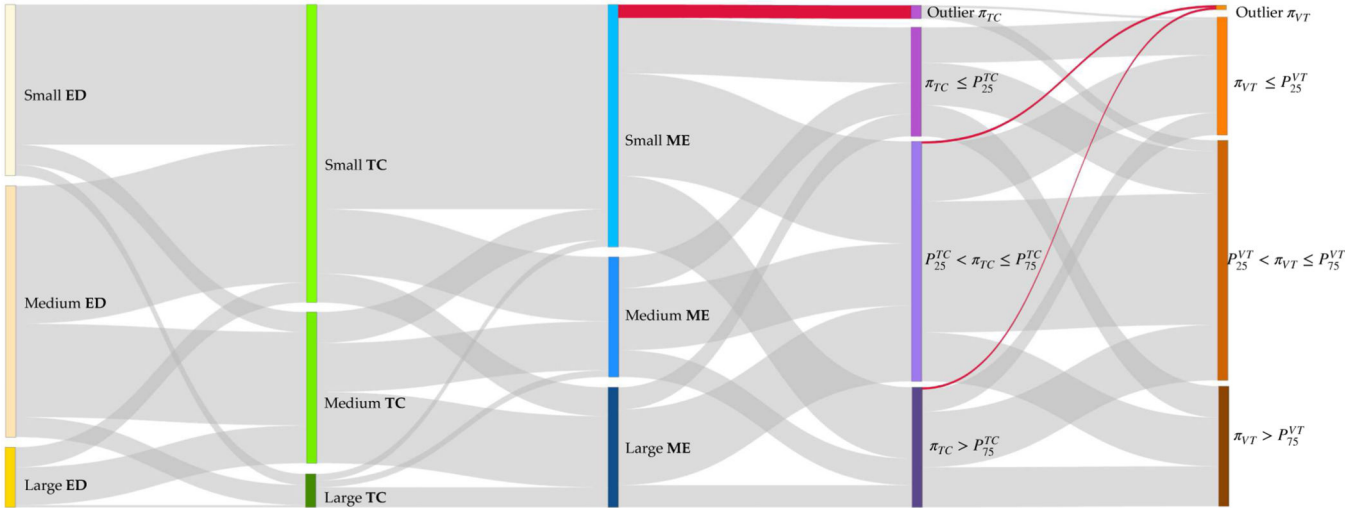
Fig. A.3:

Alluvial plot of peritumoral edema (ED) volume, tumor (TC) volume , amount of mass effect (ME), TC dice coefficients (π_{TC}) and ventricles (VT) dice coefficient (π_{VT}). We denote the ED volume less than 50 cm^3, in the range of 50 to 150 cm^3 and above 150 cm^3 to be small, medium and large ED, respectively. We use TC volume less than 50 cm^3, in the range of 50 to 100 cm^3 and above 100 cm^3 to be small, medium and large TC, respectively. We use the median and third quartile of displacement norms ($\| u \|_{\infty}$) to classify the patients as small, medium and large ME. We refer the dice coefficients less than the lower whisker limit of TC dice coefficients as outlier π_{TC} and we use first quartile (P_{25}^{TC}) and third quartile (P_{75}^{TC}) of TC dice coefficients for stratification. Similarly, we denote the lower whisker limit of VT dice coefficients as outlier π_{VT} and we use first quartile (P_{25}^{VT}) and third quartile (P_{75}^{VT}) of VT dice coefficients for classification. We highlight the links resulting to the outlier regions with red. We observe small number of subjects with small ME result in low TC dice coefficients.

REFERENCES

[1]. Steed TC, Treiber JM, Brandel MG, Patel KS, Dale AM, Carter BS, and Chen CC, "Quantification of glioblastoma mass effect by lateral ventricle displacement," Scientific Reports, vol. 8, no. 1, 2018.

[2]. Prasanna P, Mitra J, Beig N, Nayate A, Patel J, Ghose S, Thawani R, Partovi S, Madabhushi A, and Tiwari P, "Mass Effect Deformation Heterogeneity (MEDH) on Gadolinium-contrast T1-weighted MRI is associated with decreased survival in patients with right cerebral hemisphere Glioblastoma: A feasibility study," Scientific Reports, vol. 9, no. 1, pp. 1–13, Feb. 2019. [PubMed: 30626917]

[3]. Baris MM, Celik AO, Gezer NS, and Ada E, "Role of mass effect, tumor volume and peritumoral edema volume in the differential diagnosis of primary brain tumor and metastasis," Clinical Neurology and Neurosurgery, vol. 148, pp. 67–71, 2016. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0303846716302475 [PubMed: 27428485]

[4]. Subramanian S, Scheufele K, Himthani N, and Biros G, "Multiatlas calibration of biophysical brain tumor growth models with mass effect," in Medical Image Computing and Computer Assisted Intervention –MICCAI 2020, Martel AL, Abolmaesumi P, Stoyanov D, Mateus D, Zuluaga MA, Zhou SK, Racoceanu D, and Joskowicz L, Eds. Cham: Springer International Publishing, 2020, pp. 551–560.

[5]. Subramanian S, Scheufele K, Mehl M, and Biros G, "Where did the tumor start? an inverse solver with sparse localization for tumor growth models," Inverse Problems, no. 36, 2020.

[6]. "Zenodo open-sourced results," https://zenodo.org/record/7133594#.Y0EHaOzMKw6, 2022.

[7]. Babuska I. and Oden JT, "Verification and validation in computational enginerring and science: basic concepts," Computational Methods in Applied Mechanics and Engineering, vol. 193, pp. 4057–4066, 2004.

[8]. Mang A, Bakas S, Subramanian S, Davatzikos C, and Biros G, "Integrated biophysical modeling and image analysis: Application to neuro-oncology," Annual Review of Biomedical Engineering, vol. 22, no. 1, pp. 309–341, 2020, pMID: 32501772. [Online]. Available: 10.1146/annurev-bioeng-062117-121105

[9]. Lipkova J, Angelikopoulos P, Wu S, Alberts E, Wiestler B, Diehl C,Preibisch C, Pya T, Comps S, Hadjidoukas P, Leemput KV, Koumoutsakos P, Lowengrub J, and Menze B, "Personalized radiotherapy planning for glioma using multimodal bayesian model calibration," IEEE Transactions on Medical Imaging, vol. 38, no. 8, pp. 1875–1884, 2019. [PubMed: 30835219]

[10]. Mang A, Toma A, Schuetz TA, Becker S, Eckey T, Mohr C, Petersen D, and Buzug TM, "Biophysical modeling of brain tumor progression: from unconditionally stable explicit time integration to an inverse problem with parabolic PDE constraints for model calibration," Medical Physics, vol. 39, no. 7, pp. 4444–4459, 2012. [PubMed: 22830777]

[11]. Hogea C, Davatzikos C, and Biros G, "An image-driven parameter estimation problem for a reaction-diffusion glioma growth model with mass effect," J Math Biol, vol. 56, pp. 793–825, 2008. [PubMed: 18026731]

[12]. Gholami A, Mang A, and Biros G, "An inverse problem formulation for parameter estimation of a reaction-diffusion model of low grade gliomas," Journal of Mathematical Biology, vol. 72, no. 1, pp. 409–433, 2016. [PubMed: 25963601]

[13]. Swanson K, Alvord E, and Murray J, "A quantitative model for differential motility of gliomas in grey and white matter," Cell Proliferation, vol. 33, no. 5, pp. 317–330, 2000. [PubMed: 11063134]

[14]. Konukoglu E, Clatz O, Menze BH, Stieltjes B, Weber M-A, Mandonnet E, Delingette H, and Ayache N, "Image guided personalization of reaction-diffusion type tumor growth models using modified anisotropic Eikonal equations," Medical Imaging, IEEE Transactions on, vol. 29, no. 1, pp. 77–95, 2010.

[15]. Hogea C, Davatzikos C, and Biros G, "Modeling glioma growth and mass effect in 3D MR images of the brain," in Medical Image Computing and Computer-Assisted Intervention–MICCAI 2007. Springer, 2007, pp. 642–650.

[16]. Hormuth DA, Eldridge SL, Weis JA, Miga MI, and Yankeelov TE, "Mechanically coupled reaction-diffusion model to predict glioma growth: Methodological details," pp. 225–241, 2018.

[17]. Gooya A, Pohl KM, Bilello M, Cirillo L, Biros G, Melhem ER, and Davatzikos C, "GLISTR: Glioma image segmentation and registration," Medical Imaging, IEEE Transactions on, vol. 31, no. 10, pp. 1941–1954, 2013.

[18]. Scheufele K, Subramanian S, and Biros G, "Fully automatic calibration of tumor-growth models using a single mpmri scan," IEEE Transactions on Medical Imaging, pp. 1–1, 2020. [PubMed: 31135355]

[19]. Abler D, Buchler P, and Rockne RC, "Towards model-based characterization of biomechanical tumor growth phenotypes," in Mathematical and Computational Oncology, Bebis G, Benos T, Chen K, Jahn K, and Lima E, Eds. Cham: Springer International Publishing, 2019, pp. 75–86.

[20]. Swanson KR, Rockne RC, Claridge J, Chaplain MA, Alvord EC, and Anderson AR, "Quantifying the role of angiogenesis in malignant progression of gliomas: In silico modeling integrates imaging and histology," Cancer Res, vol. 71, no. 24, pp. 7366–7375, 2011. [PubMed: 21900399]

[21]. Saut O, Lagaert J-B, Colin T, and Fathallah-Shaykh HM, "A multilayer grow-or-go model for GBM: effects of invasive cells and anti-angiogenesis on growth," Bulletin of mathematical biology, vol. 76, no. 9, pp. 2306–2333, 2014. [PubMed: 25149139]

[22]. Subramanian S, Gholami A, and Biros G, "Simulation of glioblastoma growth using a 3D multispecies tumor model with mass effect," Journal of Mathematical Biology, vol. 79, no. 3, pp. 941–967, 2019. [PubMed: 31127329]

[23]. Giese A, Kluwe L, Laube B, Meissner H, Berens ME, and Westphal M, "Migration of human glioma cells on myelin," Neurosurgery, vol. 38, no. 4, pp. 755–764, 1996. [PubMed: 8692396]

[24]. Lemée J-M, Clavreul A, and Menei P, "Intratumoral heterogeneity i   glioblastoma: don't forget the peritumoral brain zone," Neuro-oncology, vol. 17, no. 10, pp. 1322–1332, 2015. [PubMed: 26203067]

[25]. Myronenko A, "3d mri brain tumor segmentation using autoencoder regularization," 2018.

[26]. Bakas S, Reyes M, and et al. , "Identifying the best machine learning algorithms for brain tumor segmentation, progression assessment, and overall survival prediction in the BRATS challenge," CoRR, vol. abs/181102629, 2018. [Online]. Available: http://arxiv.org/abs/1811.02629

[27]. Avants BB, Tustison N, and Song G, "Advanced normalization tools (ants)," Insight j, vol. 2, no. 365, pp. 1–35, 2009.

[28]. Gholami A, Mang A, Scheufele K, Davatzikos C, Mehl M, and Biros G, "A framework for scalable biophysics-based image analysis," in Proc ACM/IEEE Conference on Supercomputing, no. 19, 2017, pp. 19:1–19:13.

[29]. Kofler F, Paetzold JC, Ezhov I, Shit S, Krahulec D, Kirschke JS, Zimmer C, Wiestler B, and Menze BH, "A baseline for predicting glioblastoma patient survival time with classical statistical models and primitive features ignoring image information," in Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries, Crimi A. and Bakas S, Eds. Cham: Springer International Publishing, 2020, pp. 254–261.

[30]. McKinley R, Rebsamen M, Daetwyler K, Meier R, Radojewski P, and Wiest R, "Uncertainty-driven refinement of tumor-core segmentation using 3d-to-2d networks with label uncertainty," 2020.

[31]. McKinley R, Rebsamen M, Daetwyler K, Meier R, Radojewski P, and Wiest R, "Uncertainty-driven refinement of tumor-core segmentation using 3d-to-2d networks with label uncertainty," 2020.

[32]. Bommineni V, "Piecenet: A redundant unet ensemble," 2020.

[33]. Ali MJ, Akram MT, Saleem H, Raza B, and Shahid AR, "Glioma segmentation using ensemble of 2d/3d u-nets and survival prediction using multiple features fusion," in Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries, Crimi A. and Bakas S, Eds. Cham: Springer International Publishing, 2021, pp. 189–199.
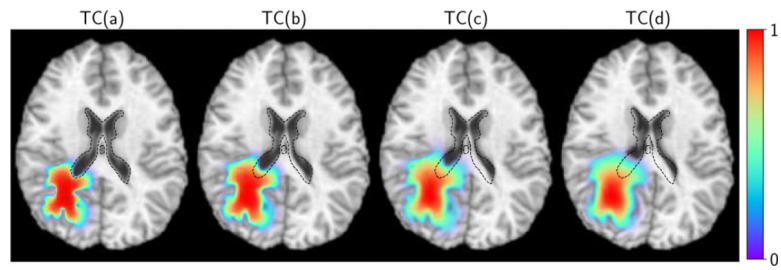
Fig. 1:
Synthetic data T1 MRIs generated with our forward model. The tumor concentration is overlaid (color) along with the undeformed ventricles (black dashed line) to indicate the variable extent of mass effect. The maximum displacements are approximately 0, 5.4, 8.9, and 13.3 mm, respectively.
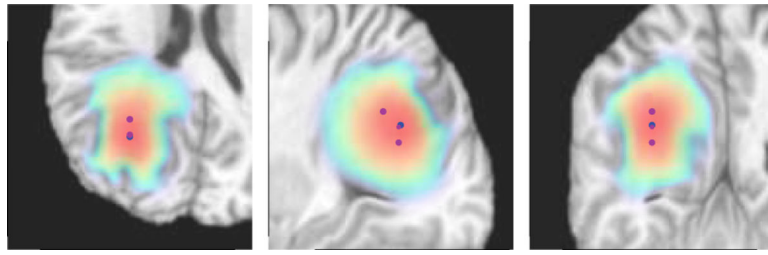
Fig. 2:

Estimated initial conditions p reconstructed using the algorithm in (**S**.1) for test-case TC(d). The size of the markers indicate the magnitude of activation. Further, the initiation locations are projected onto the respective 2D slice view (axial, sagittal, coronal) for ease of visualization. The blue marker is the ground truth and the magenta markers represent the reconstruction. The tumor concentration is overlaid (red: high, green: low). Due to ill-posedness, the exact reconstructions cannot be recovered.
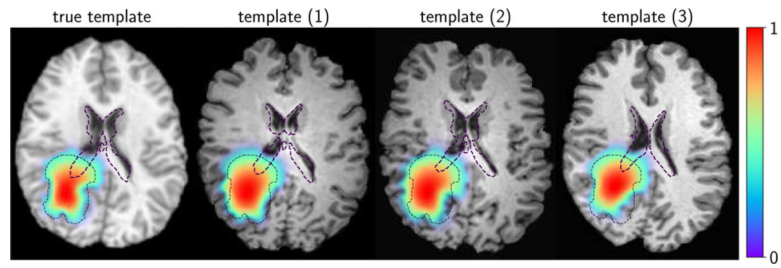
Fig. 3:

Reconstruction of the patient using 16 templates (of which three templates are shown here) as the precancer scan. The tumor data segmentation is highlighted as a black dashed contour line. The reconstructed tumor concentration is shown in color. The undeformed ventricle configuration is highlighted as a purple dashed line to indicate the extent of mass effect in the reconstructions.
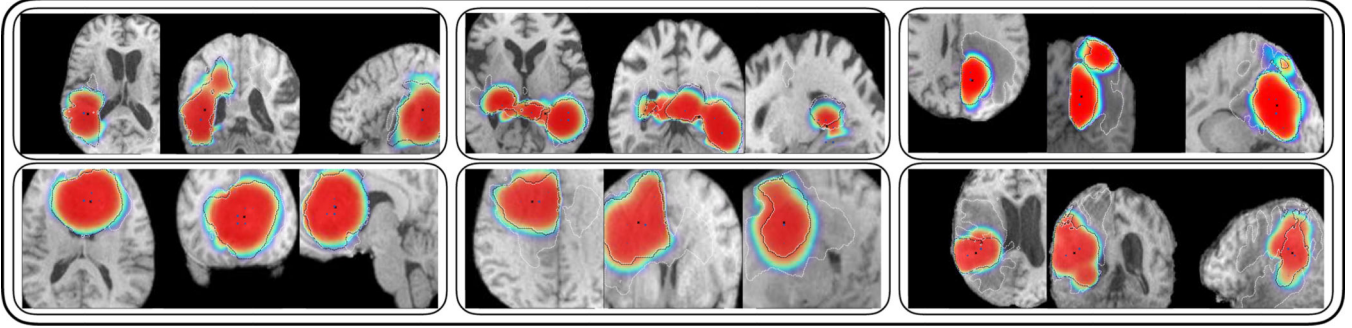
Fig. 4:

Estimated initial conditions p reconstructed using the algorithm in ($\mathbf{S}$.1) for a few subjects. Each patient is a set of three images representing the axial, sagittal, and coronal 2D slices of the patient brain, zoomed into the tumor region. The blue markers represent the tumor initiation reconstruction and the black cross is the center of mass of the tumor core component. The size of the markers indicate the magnitude of activation. Further, the initiation locations are projected onto the respective 2D slice view (axial, sagittal, coronal) for ease of visualization. The tumor concentration is overlaid (red: high, green: low) with the black dashed line representing the tumor core data and the white dashed line representing edema.
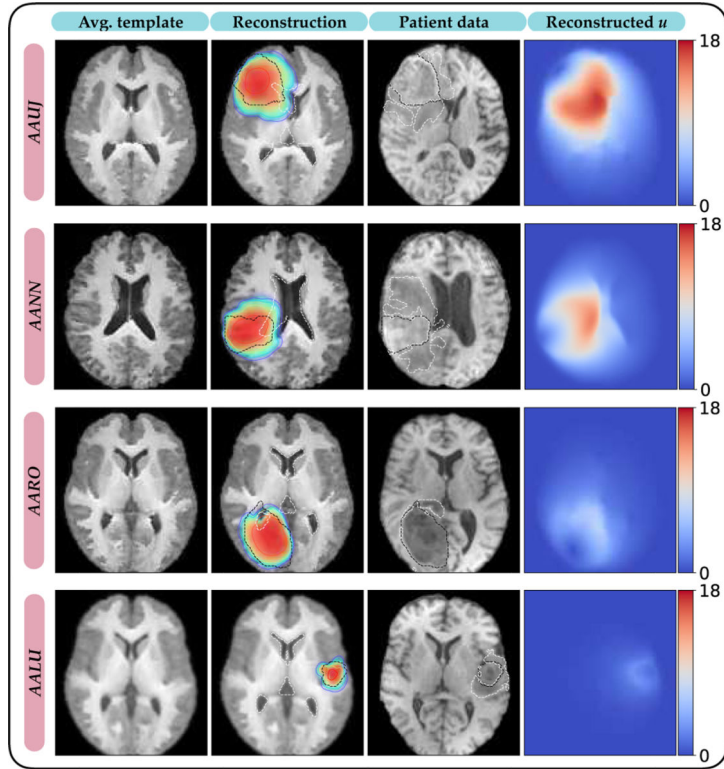
Fig. 5:
Reconstructed images from the inversion scheme. We show four representative subjects (IDs in parenthesis). The first column shows the averaged template T1 MRI, the second column depicts the averaged patient reconstruction with the estimated tumor concentration (higher concentrations (~1) are indicated in red and lower ones in green) overlayed. We also show the undeformed ventricle contour (white dashed line) to express the extent of mass effect, as well as the TC data contour (black dashed line). The third column shows the patient T1 MRI data with the TC data contour (black dashed line) and the edema data contour (white dashed line). The last column shows the reconstructed average displacement (in mm) two-norm field.
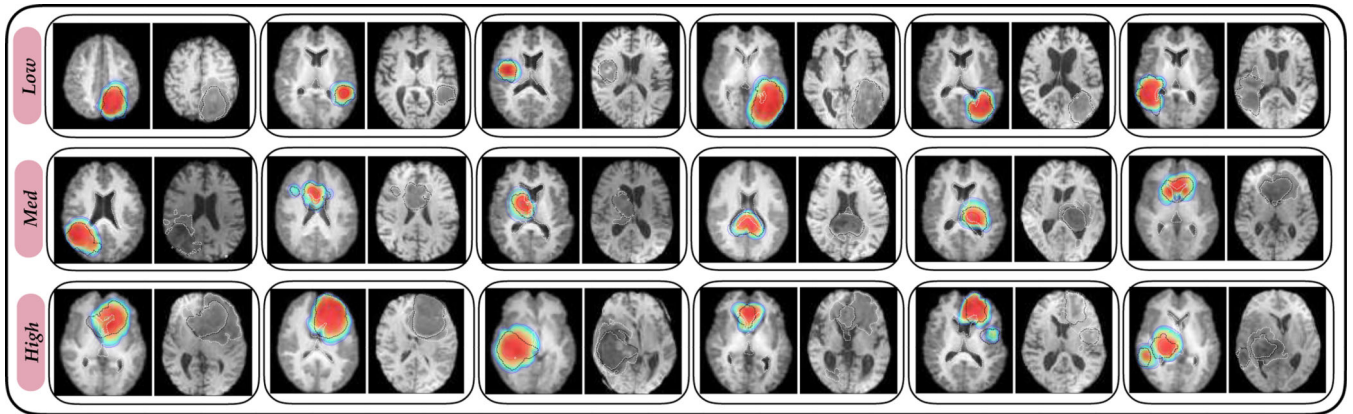
Fig. 6:
Some clinical reconstructions stratified into low, medium, and high mass effect based on equal quantiles of displacement. Each image is a pair indicating the reconstructed patient with estimated tumor concentrations (left) and patient T1 MRI data (right). The TC data (black dashed line) and undeformed ventricles (white dashed line) are visualized in the reconstructed patient (left) and TC data (black dashed line) and edema data (white dashed line) are visualized in the patient data (right).

(a)

(b)
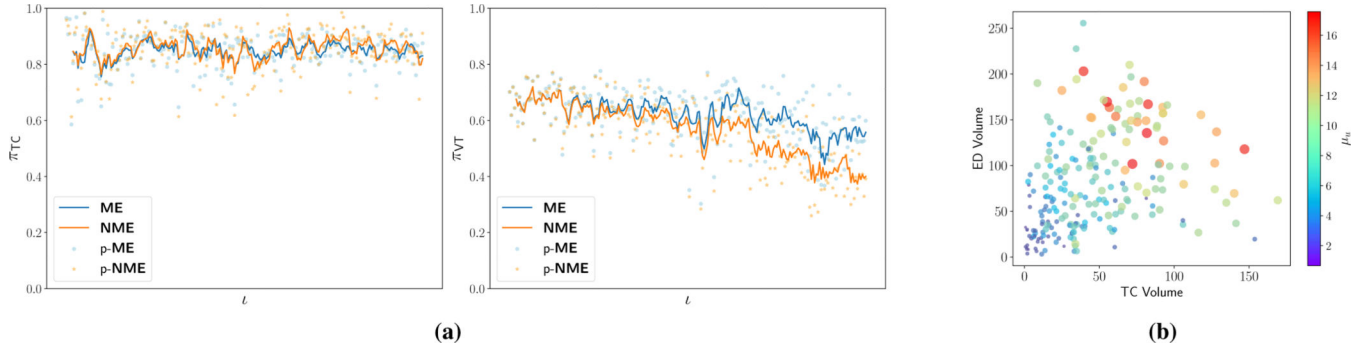
Fig. 7:

(a) **CQ1**: Plots of tumor core (TC, left) dice coefficient π_{TC} and ventricle (VT, right) dice coefficient π_{VT} vs patient ID index ι sorted according to increasing mass effect for the two models **ME** and **NME**. The lines (**ME,NME**) represent a rolling 5-patient mean and p-**ME** and p-**NME** are the raw data points. We observe similar reconstruction for TC, but significant performance gains are visible in the VT reconstruction. (b) **CQ2**: Scatter plot of tumor core (TC) volume vs peritumoral edema (ED) volume (volumes in cm^3). Each data point is sized and colored according to the mean displacement field μ_u—larger (and red) circles are indicative of patients with high mass effect. We observe variability in mass effect deformations across tumors with similar volume. For example, tumors with TC volume approximately 75 cm^3 show a wide range of mass effect.
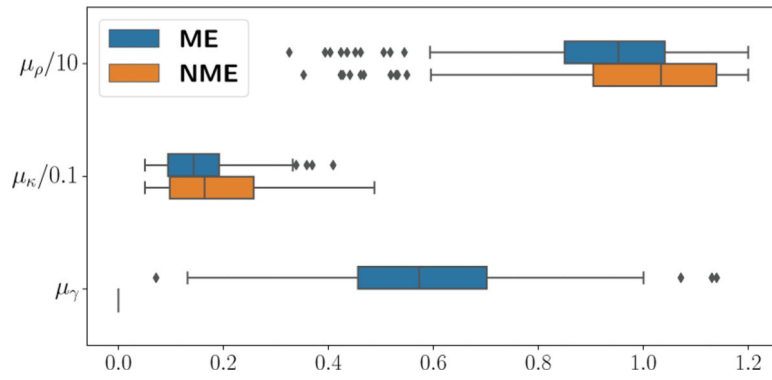
Fig. 8:

Box plots of the mean reaction, diffusion, and mass effect coefficients ($\mu_\rho, \mu_\kappa, \mu_\gamma$, respectively) for the patient cohort using the two models **ME** and **NME**. We observe mass effect in the range $\mu_u = 7.18 \pm 3.93$ mm (μ_u is the mean of $u_\infty = \| u \|_\infty$, where u is the displacement two-norm field).
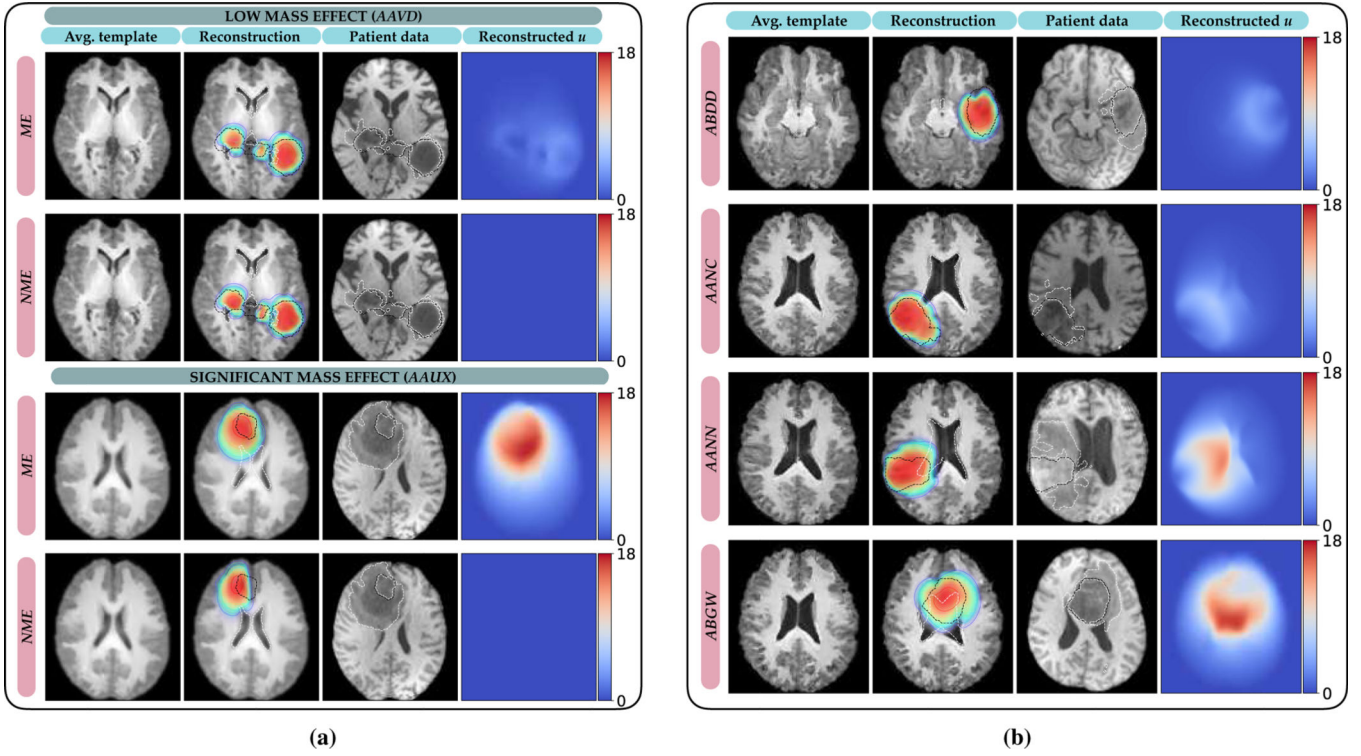
Fig. 9:

Reconstructed images from the inversion scheme. (**a**) We compare the two models **ME** and **NME**. We compare the results for one subject with low and one subject with high mass effect. The first column shows the averaged template T1 MRI, the second column shows the averaged patient reconstruction with the estimated tumor concentration (higher concentrations (~1) are indicated in red and lower ones in green) overlayyed. We also show the undeformed ventricle contour (white dashed line) to express the extent of mass effect, as well as the TC data contour (black dashed line). The third column shows the patient T1 MRI data with the TC data contour (black dashed line) and the edema data contour (white dashed line). The last column depicts the reconstructed average displacement (in mm) two-norm field. Both models exhibit similar reconstruction performance for the tumor and healthy tissues in patient AAVD (low mass effect); for patient AAUX (high mass effect) **ME** outperforms **NME** because it is capable of deforming the ventricles (and other structures; for example, the model captures the mid-line shift necessary to account for the tumor). (**b**) Reconstructions for patients with similar TC volume but different ranges of mass effect. Each row represents a subject (with subject ID indicated on the left). We observe different mass effect ranges ($\mu_u = 4.1, 6.9, 11.7, 17.3$ mm from top to bottom) for similar TC volumes ($65.6, 69.0, 70.1, 72.0$ cm^3 from top to bottom). While the first patient shows almost no mass effect, the last patient exhibits significant compression of ventricles (also seen in the deformation from the initial ventricle contours and the displacement field).
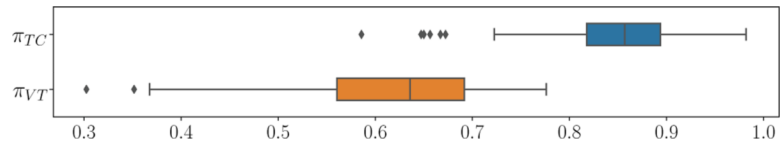
Fig 10:

Box plot of tumor core (TC) dice coefficients (π_{TC}) and ventricle (VT) dice coefficients (π_{VT}) for the patients cohort using **ME** model. We observe π_{TC} in the range of 0.586 to 0.982 and π_{VT} in the range of 0.302 to 0.776 (See Fig. A.3).
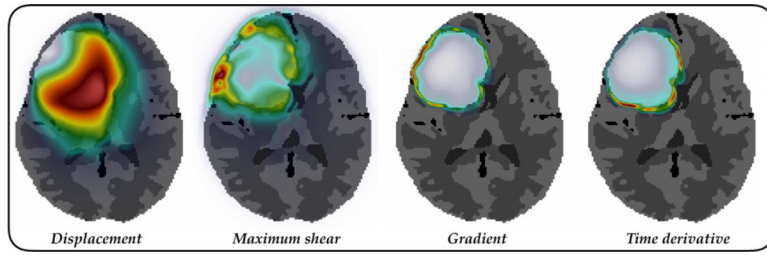
Fig. 11:

Voxelwise biophysics-based features for a representative patient. We plot 2D slices of $\| \boldsymbol{u} \|_2 , \tau, \| \nabla c \|_2 , \partial c / \partial t$ at time instant when the tumor volume fraction is 5% of the brain volume. All feature values are normalized here (high: red, low: blue)
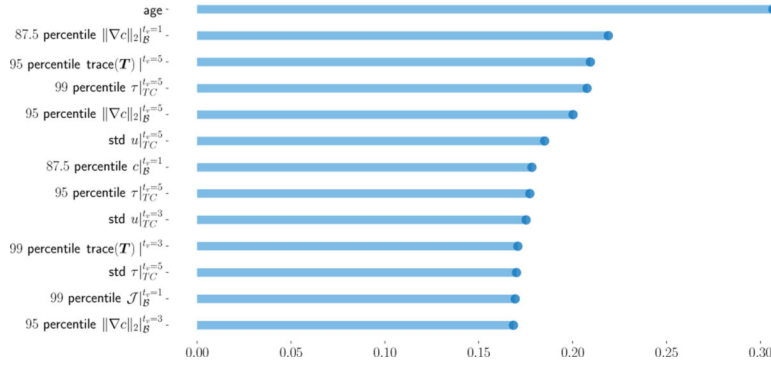
Fig. 12:

Spearman correlation of features (only statistically significant correlations, $p < 0.05$) with survival sorted in descending order of values. For any feature statistic y, $y|_{\mathcal{R}}^{t_v = v}$ represents the statistic computed in region $\mathcal{R} \in \{\mathcal{B}, TC\}$ (absent $\mathcal{R}$ denotes no region and statistic is computed using all x) at time when tumor volume fraction t_v is v.
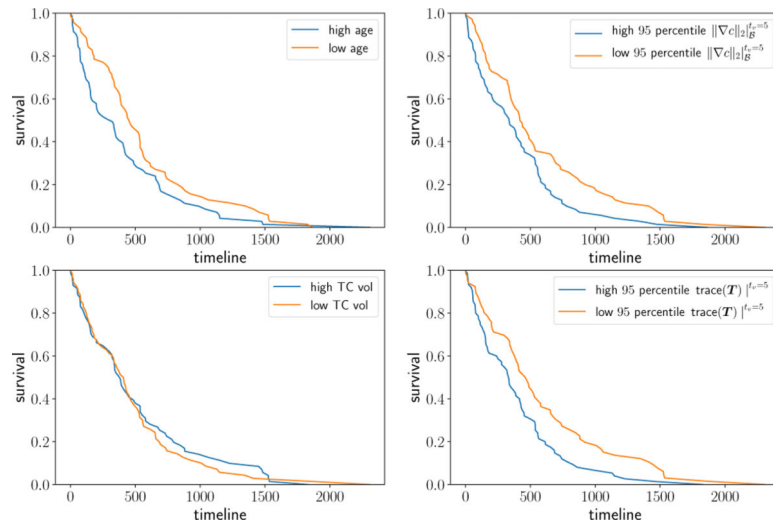
Fig. 13:

Kaplan-Meier survival curves using median feature value patient stratification for some top correlated (with survival) features and tumor core volume. Age (range of 24 to 88 years, median 65 years) and biophysics feature survival curves are significant (log rank p < 0.05). For tumor core volume (range of about $1cm^3$ to $150cm^3$ with median of $40cm^3$), log rank p = 0.72 and is insignificant—subjects are not stratified based on volumetric features.
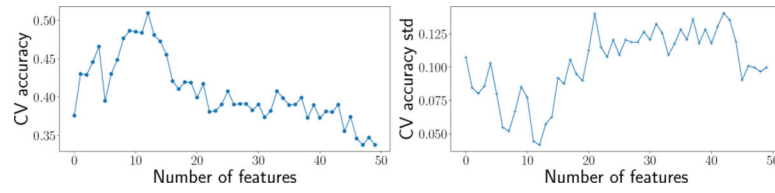
Fig. 14:
CV accuracy and standard deviation of accuracy values of a random forest classifier by sequentially adding features in descending order of their feature importances. We select the top 13 features due to their maximum CV accuracy and small accuracy std values.

TABLE I:

Inversion results assuming **known** m_0, the anatomy of the subject before tumor occurrence. We report our reconstructions for parameters $\iota = \gamma, \rho, \kappa$, and the maximum value of the displacement field two-norm (i.e., $u_\infty = \parallel u \parallel_\infty$). The ground truth parameters are indicated by $\iota^\star$ and the reconstructed values are ι. We also report e_ι, the relative error for parameters $\iota = \gamma, \rho, \kappa$; e_u is the relative error in norm of the displacement field two-norm (i.e., $\parallel u \parallel_2$). Finally, we report π_{TC}, the dice coefficient in reconstructing the tumor core. The tumor initial condition used for inversion is indicated in the column "Tumor IC" – True IC indicates synthetic ground truth IC and Inverted IC indicates our reconstructions using (**S**.1) and (**S**.3). (Note: when the ground truth ($\star$) is zero, the errors are absolute errors; diffusion coefficient (κ) values are scaled by 1E–2 for clarity).

Test-case	Tumor IC	$(\gamma^\star, \gamma)$	e_γ	$(\rho^\star, \rho)$	e_ρ	$(\kappa^\star, \kappa)/0.01$	e_κ	$(u_\infty^\star, u_\infty)$	e_u	π_{TC}
TC (a)	True IC	(0,0.0053)	5.31E–3	(12,12.01)	9.58E–4	(2.5,2.48)	8.23E–3	(0,0.07)	1.92E–3	9.96E–1
TC (b)	True IC	(0.4,0.4)	1.58E–3	(12,11.99)	8.33E–5	(2.5,2.49)	1.41E–3	(5.4,5.43)	6.87E–3	9.96E–1
TC (c)	True IC	(0.8,0.79)	4.68E–4	(10,10)	6.30E–5	(5,5)	2E–5	(8.98,8.97)	1.12E–3	9.98E–1
TC (d)	True IC	(1.2,1.19)	6.67E–5	(10,9.99)	1E–5	(2.5,2.5)	1.46E–3	(13.35,13.37)	3.36E–3	9.97E–1
TC (a)	Inverted IC	(0,0.082)	8.22E–2	(12,9.82)	1.81E–1	(2.5,2.41)	3.41E–2	(0,0.11)	8.25E–1	9.06E–1
TC (b)	Inverted IC	(0.4,0.42)	2.99E–2	(12,10.13)	1.55E–1	(2.5,2.02)	1.93E–1	(5.4,5.18)	2.17E–2	9.03E–1
TC (c)	Inverted IC	(0.8,0.82)	2.36E–2	(10,8.62)	1.38E–1	(5,4.17)	1.67E–1	(8.98,8.55)	3.54E–2	9.98E–1
TC (d)	Inverted IC	(1.2,1.18)	1.73E–2	(10,8.5)	1.49E–1	(2.5,1.73)	3.05E–1	(13.35,12.25)	7.68E–2	8.73E–1

TABLE II:

Inversion results for **unknown m_0** and **unknown p** using the multistage inversion scheme (see §2-C). We report the ground truth parameters $\iota^\star$ and the mean (μ_ι) and standard deviation (σ_ι) of the reconstructed parameters across the ensemble of templates; $\iota = \gamma, \rho, \kappa$, and the maximum value of the displacement field two-norm (i.e., $\| u \|_\infty$ in mm). e_ι is the relative error for the average value of biophysical parameter $\iota = \gamma, \rho, \kappa$; e_u is the relative error in the average norm of the displacement field two-norm (i.e., $\| u \|_2$), and π_{TC} is the probabilistic tumor core dice coefficient. (Note: when the ground truth ($\star$) is zero, the errors are absolute errors; diffusion coefficient (κ) values are scaled by $1E-2$ for clarity).

Test-case	$(\gamma^\star, \mu_\gamma, \sigma_\gamma)$	e_γ	$(\rho^\star, \mu_\rho, \sigma_\rho)$	e_ρ	$(\kappa^\star, \mu_\kappa, \sigma_\kappa)/0.01$	e_κ	$(u_\infty^\star, \mu_{u_\infty}, \sigma_{u_\infty})$	e_u	π_{TC}
TC (a)	(0.0,0.19,0.15)	1.97E−1	(12.8,96,0.65)	2.53E−1	(2.5,1.89,0.74)	2.41E−1	(0.2,1.1,1.51)	1.76E1	7.93E−1
TC (b)	(0.4,0.43,0.2)	7.93E−2	(12.8,82,0.32)	2.65E−1	(2.5,2.41,0.61)	3.34E−2	(5.4,4.93,2.21)	8.48E−2	8.20E−1
TC (c)	(0.8,0.69,0.2)	1.25E−1	(10,8,0.15)	2E−1	(5,4.23,0.83)	2.41E−1	(8.98,6.89,1.86)	2.22E−1	8.48E−1
TC (d)	(1.2,1.1,0.14)	7.76E−2	(10,7.78,0.26)	2.23E−1	(2.5,2.69,0.72)	7.67E−2	(13.35,10.43,1.09)	2.02E−1	8.39E−1

TABLE III:

Inversion results for clinical data. We report the mean (μ_ι) and standard deviation (σ_ι) of the reconstructed parameters across the ensemble of templates; $\iota = \gamma, \rho, \kappa$, and the maximum value of the displacement field two-norm (i.e., $u_\infty = \| u \|_\infty$ in mm). We also report the probabilistic tumor core dice coefficient π_{TC}, the average runtime μ_T, and standard deviation in runtime σ_T in (hh:mm:ss). (Note: the diffusion coefficient (κ) values are scaled by 1E − 2 for clarity).

Patient ID	$(\mu_\gamma, \sigma_\gamma)$	(μ_ρ, σ_ρ)	$(\mu_\kappa, \sigma_\kappa)/0.01$	$(\mu_{u_\infty}, \sigma_{u_\infty})$	π_{TC}	(μ_T, σ_T) (in hh:mm:ss)
AAUJ	(0.82, 0.081)	(9.92, 0.75)	(1.87, 0.89)	(17.57, 1.4)	0.875	(01:13:27, 00:13:43)
AANN	(0.59, 0.25)	(9.93, 0.61)	(2.28, 1.23)	(11.68, 4.91)	0.853	(01:27:24, 00:06:23)
AARO	(0.4, 0.089)	(10.64, 0.58)	(1.75, 0.38)	(7.23, 1.49)	0.839	(01:08:19, 00:05:47)
AALU	(0.36, 0.23)	(6.39, 0.52)	(1.1, 0.55)	(2.43, 1.48)	0.942	(01:05:22, 00:08:14)

TABLE IV:

Classification of patient survival using age, volumetric, and biophysics-based features. We show the improvement in accuracy by using the features subselected by random forest feature importance—these features include age, volumetric, and biophysics-based features.

Features	CV training accuracy	Testing accuracy
Age	0.375 ± 0.107	0.276
Age + Volumetric	0.347 ± 0.073	0.345
Selected features	0.509 ± 0.042	0.448

TABLE V:

Classification of patient survival using age, volumetric, and biophysics-based features on the BraTS 2020 dataset. We repeat our inversion on the BraTS dataset and report training cross-validated accuracies, testing accuracy on a random 20% split, and the accuracy on the leaderboard validation dataset. For the learderboard accuracies, we compare with the top three winning algorithms [31]–[33] for the BraTS 2020 competition (bolded for comparison). We observe that biophysics-based features assist in improving accuracies (as in Tab IV) and provide competitive accuracies on the challenge dataset.

Features	CV training	Testing	Leaderboard
Age	0.358 ± 0.051	0.428	**0.379**
Age + Volumetric	0.432 ± 0.031	0.457	**0.414**
Selected features	0.567 ± 0.040	0.485	**0.483**
BraTS 1st/2nd/3rd place	-	-	**–/0.379/0.483**