ORIGINAL ARTICLE

Journal Section

Optimal Mass Transport Kinetic Modeling for Head and Neck DCE-MRI: Initial Analysis

Rena Elkin[1]* | Saad Nadeem[2]* | Eve LoCastro[2] | Ramesh Paudyal[2] | Vaios Hatzoglou[3] | Nancy Y. Lee[4] | Amita Shukla-Dave[2,3] | Joseph O. Deasy[2] | Allen Tannenbaum[5]

[1]Stony Brook University, Applied Mathematics & Statistics, Stony Brook, NY 11794, USA

[2]Memorial Sloan Kettering Cancer Center, Department of Medical Physics, New York, NY 10065, USA

[3]Memorial Sloan Kettering Cancer Center, Department of Radiology, New York, NY 10065, USA

[4]Memorial Sloan Kettering Cancer Center, Department of Radiation Oncology, New York, NY 10065, USA

[5]Stony Brook University, Computer Science and Applied Mathematics & Statistics, Stony Brook, NY 11794, USA

Correspondence
Saad Nadeem, Department of Medical Physics, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA
Email: nadeems@mskcc.org

Funding information
AFOSR, Grant Number: FA9550-17-1-0435; National Institutes of Health, Grant Number: R01-AG048769 and R01-CA198121; MSK Cancer Center Support Grant, Grant Number: P30 CA008748; BRCF, Grant Number: BCRF-17-193

Abstract: Dynamic contrast-enhanced (DCE) magnetic resonance imaging (MRI) provides a quantitative measure of pharmacokinetic (PK) parameters in body tissues; it entails collection of a series of T1-weighted images following the administration of a paramagnetic contrast agent (CA). Current state-of-the-art models for estimating the PK parameters do not account for intervoxel movement of the contrast agent. We introduce an optimal mass transport (OMT) formulation that naturally handles intervoxel CA movement and distinguishes between advective and diffusive flows. Ten patients with head and neck squamous cell carcinoma (HNSCC) were enrolled in the study between June 2014 and October 2015 and underwent DCE MRI imaging prior to beginning treatment. Converting DCE image intensity to CA tissue concentration, this information was taken as the input in the OMT model, generating estimates of forward flux (Φ_F) and backward flux (Φ_B). We decompose computed flux into advective and diffusive flows which are known to occur due to the typical pressure gradients on the tumor boundary. We additionally show that we can capture the reflux with our backward flux (Φ_B). We extensively test our approach on HNSCC data and provide additional quantitative information such as Φ_F, Φ_B, advective and diffusive flows. The approach given in the present paper is essentially data driven in which a novel optical flow scheme is applied to the imagery in order to analyze the flow.

KEYWORDS

Optimal Mass Transport, DCE-MRI, Pharmacokinetic Modeling, Advection, Diffusion

* Equally contributing authors.

1 | INTRODUCTION

Head and neck (HN) tumors are heterogeneous with complex anatomy [1]. Accurate detection and delineation of tumor extent is critical to optimize treatment planning; patients therefore routinely undergo non-invasive imaging for careful assessment of this complex anatomy. Non-invasive magnetic resonance imaging (MRI) has served an important role as a diagnostic test for initial staging and follow-up tumors in the HN region [2, 3, 4].

Dynamic contrast-enhanced (DCE) MRI involves the acquisition of successive T1-weighted images with administration of a T1-shortening Gadolinium-based contrast agent (CA) [5]. DCE imaging has been considered to be a promising tool for clinical diagnostics, including head and neck cancers [6, 7]. DCE imaging data can be evaluated by semi-quantitative analysis and model based quantitative analysis [5, 8].

The Tofts-Kermode pharmacokinetic model is the most commonly used model for DCE-MRI analysis [5]. In DCE-MRI, CA concentration is derived from changes in signal intensity over time, then fitted to tracer model to estimate pharmacokinetic (PK) parameters related to tumor vascular permeability and tissue perfusion [5]. The volume transfer constant for CA (K^{trans}) for contrast agent from the blood plasma into the extravascular extracellular space (EES).

The accuracy and precision of this quantitative parameter can be influenced by arterial input function (AIF) quantification, temporal resolution in data acquisition, signal-to-noise ratio (SNR), and pharmacokinetic model selection [9, 10, 11, 12]. Moreover these quantitative metrics are macroscopic, representing averages over highly heterogeneous microscopic processes.

In the present work, we propose to employ methods from optimal mass transport (OMT) theory to study the problem of estimating CA flow from tumor DCE-MRI data. The problem of OMT dates back to Monge in 1781, who posed the problem of finding the minimal transportation cost for moving a pile of soil from one location to another. OMT was given a modern and relaxed formulation in the work of Kantorovich, and so is now known as the Monge–Kantorovich problem. Applications include image processing and computer vision, econometrics, fluid flows, statistical physics, machine learning, expert systems, and meteorology [13, 14].

The dynamical version of OMT, put forth by Benamou and Brenier [15], opened up new possibilities for numerical solutions and extensions using tools from the field of fluid dynamics. Their approach considers the minimization of kinetic energy subject to a continuity equation that enforces mass preservation. For the approach here, we add a diffusion term to the continuity equation. The resulting advection/diffusion equation regularizes the flow and moreover, as we will argue, gives a better physical picture of the underlying dynamics. Indeed, it has been demonstrated that the differences in elevated tumor pressure lead to exudate convective flow out of the tumor such that any flow into the tumor may be ascribed to diffusion [16]. It is important to emphasize that our analysis is data driven, and all of our conclusions are based on the imagery.

Other than using the advection/diffusion flow from computational fluid dynamics, we do not impose any prior constraints or assumptions of the underlying physics of the model; for example, we do not require AIF as an input. This type of approach was exploited in our previous work [17] for the study of the glymphatic system. We should note, if one integrates the advection-diffusion equation over a given “compartment” and applies OMT theory, one derives an equation very similar to the Tofts-Kermode model [18].

2 | METHODS

2.1 | Optimal Mass Transport with diffusion

The main tool we will be employing is based on the theory of optimal mass transport (OMT). We will not give all the details here, but simply refer the interested reader to [13, 14]. The theoretical details relevant to this work are provided in the supplementary material; some of the equations in the supplementary are referred here for completeness.

In this work, we modify the original Benamou and Brenier OMT formulation by adding a diffusion term in the continuity equation to allow for ***both advection and diffusion*** of the contrast agent and having the added benefit of regularizing the flow. Various aspects of regularized OMT have been studied in a number of recent works including [19, 20, 21, 22], which have extensive lists of references.
Accordingly, we formulate the following ***regularized OMT*** problem

$$\inf_{\mu,v} \int_0^1 \int_A \quad \frac{1}{2}\mu(t,x)\|v(t,x)\|^2 \, dx \, dt \tag{1}$$

$$\text{subject to} \quad \partial_t \mu + \nabla \cdot (\mu v) = \nabla \cdot (D \nabla \mu), \tag{2}$$

$$\mu(0,\cdot) = \mu_0(\cdot), \quad \mu(1,\cdot) = \mu_1(\cdot), \tag{3}$$

where D, encapsulating spatially-varying diffusion, is a positive definite symmetric matrix modelling the diffusion. where $\mu = \mu(t,x) \in \mathbb{R}$ is the density interpolant between the given densities μ_0 and μ_1, assumed to have the same total mass, $v = v(t,x) \in \mathbb{R}^d$ is the velocity, and $A \subset \mathbb{R}^d$ is a subdomain. For a more complete derivation of the expressions comprising OMT, we refer the reader to the appended Supplemental Theory section.

2.1.1 | OMT relation to Tracer Kinetic Model

We very briefly sketch the relation to the very popular Tofts-Kermode model [5]. For simplicity, we indicate details in 2D. The basic idea is to integrate the advection-diffusion equation (2) over a given compartment V and use the divergence theorem. Accordingly, we let S denote the boundary of V. We first note that

$$\int_V \nabla \cdot (\mu v) \, dV = \int_S \mu v \cdot \vec{\mathbf{n}} \, dS \tag{4}$$

$$= -\int_S \mu \nabla \lambda \cdot \vec{\mathbf{n}} \, dS =: \mathcal{P}_\mu. \tag{5}$$

where $\vec{\mathbf{n}}$ denotes the outward normal to the surface S.

This part, derived from advection, may be interpreted as arising from Darcy's law (difference in pressure) but because of (22), there is also the influence of diffusion. We should emphasize that OMT theory gives a unique solution of (2) with optimal velocity given as the gradient of a function via (22).
We see that at the optimal solution of the OMT problem, (2) becomes an ordinary differential equation which leads to pharmacokinetic modelling. The rate of change of CA density over the compartment can be written as

$$\frac{d\rho(t)}{dt} = \int_S (D\nabla\mu - \mu v) \cdot \vec{\mathbf{n}} \, dS, \tag{6}$$

where $\vec{\mathbf{j}} = D\nabla\mu - \mu v$ is the total flux. Although the direction of the flux, $\vec{\mathbf{j}}$, is influenced by both the density gradient

$\nabla\mu$ and the velocity v, we chose to focus our attention on the velocity since it is affected by time-varying changes in the given images as well as diffusion, as seen from the Hamilton-Jacobi equation (22). Interestingly, we find that the velocity alone yields information that aligns well with the current understanding of the contrast behavior which we will elaborate on below. The full derivation is presented in the appended Supplemental Theory section.

2.2 | Full OMT model for flow of contrast agent

In this section, we summarize our overall model based on the regularized OMT model. In what follows, we will take the diffusion matrix $D = \sigma^2 I$, that is, a constant times the identity. In other words, diffusion is not a tensor quantity, and there are no preferred directions.

We make the following assumptions about the data to motivate the proposed regularized model:

(i) Image intensity is proportional to contrast agent density. Intensity is taken as the mass density in the OMT model.

(ii) The flow is governed by the advection-diffusion equation (ADE),

$$\partial_t \mu + \nabla \cdot (\mu v) = \sigma^2 \Delta \mu, \tag{7}$$

where $0 \leq \mu : [0, T] \times A \to \mathbb{R}$ is a density with compact support, $A \subset \mathbb{R}^d$ is a subdomain, $v : [0, T] \times A \to \mathbb{R}^d$ is velocity and $0 \leq \sigma \in \mathbb{R}$ is diffusivity. As alluded to above, we take $d = 3$ in what follows. Further, in our treatment above we took $T = 1$. Here we take a general final time T. The mathematical treatment is identical.

(iii) We are given noisy observations of the contrast agent density at times $t = T_i$,

$$\mu(T_i, x) + \gamma = \mu^{\text{obs}}_{T_1}(x), \quad i = 0, \ldots, N, \tag{8}$$

where γ is Gaussian iid noise with covariance Σ.

Given initial and final observations of contrast agent density μ^{obs}_0 and μ^{obs}_T at times $t = 0$ and $t = T$ respectively, we need to determine the velocity field v and the 'true' or 'clean' images μ_0 and μ_T such that the constraint (7) is satisfied. Accordingly, we propose to minimize (with respect to v and μ) the objective function

$$\int_0^T \int_A \mu \|v\|^2 \, dx dt + \epsilon^2 \|\mu(T, x) - \mu^{\text{obs}}_T(x)\|^2_\Sigma, \tag{9}$$

subject to the ADE constraint (7). The notation indicates that the variance operation is inversely scaled by the covariance terms.

Here, the second term is added to the Benamou and Brenier energy functional [15] to prevent over-fitting to noise in the data. The parameter ϵ^2 drives the balance between fitting the data and minimizing the energy associated with transporting the mass. We refer to (9) (with constraint (7)) as the ***generalized regularized OMT*** problem (GR-OMT).

2.3 | Patient Studies

The institutional review board approved and granted a waiver of informed consent for this retrospective clinical study, which was compliant with the Health Insurance Portability and Accountability Act. A total of 10 patients with neck nodal metastases and histologically-proven squamous cell carcinoma were enrolled into the study between June 2014

and October 2015.

2.4 | DCE-MRI data acquisition

MRI studies were performed on a 3-Tesla (T) scanner (Ingenia; Philips Healthcare; Netherlands) using a neurovascular phased-array coil [23]. The standard MR acquisition parameters were as follows: multiplanar (axial, coronal, and sagittal) T2-weighted (T_2w), fat-suppressed, fast spin-echo images (repetition time [TR]=4000 ms; echo time [TE]=80 ms; number of averages (NA)=2; matrix=256x256; slice thickness=5 mm; field of view [FOV]=20–24 cm), and multiplanar T1-weighted (T_1w) images (TR=600 ms; TE=8 ms; NA=2; slice thickness=5 mm; matrix=256x256; slice thickness=5 mm; FOV=20–24 cm). Multiple flip angle pre-contrast T_1w images were subsequently acquired for T_{10} measurement and DCE-MRI. Pre-contrast T_1w images were acquired using a fast-multiphase spoiled gradient recalled (SPGR) echo sequence and the acquisition parameters were as follows: TR = 7 ms, TE = 2.7 ms, flip angles $\theta = 5^\circ$, 15° , 30°, single-excitation, NA = 1, pixel bandwidth = 250 Hz/pixel, 256 x 128 matrix that was zero filled to 256 x 256 during image reconstruction and FOV = 20-24 cm^2.

The dynamic imaging sequence was acquired using the parameters for pre-contrast T_1w with a flip angle of 15° . After acquisition of the first 5-6 images, antecubital vein catheters delivered a bolus of 0.1 mmol/kg gadolinium-based clinically-approved contrast at 2 mL/s, followed by saline flush using an MR-compatible programmable power injector (Spectris; Medrad, Indianola, PA). The entire node was covered with 5-mm thick slices, zero gap, resulting in collection of 8-10 slices with 7.46-8.1 second temporal resolution, for 40-60 time course data points. The time duration of dynamic imaging ranged 5-6 minutes.

2.5 | DCE-MRI Theory

The T_1-weighted (T_1w) DCE signal for the spoiled gradient-echo recalled sequence is given by [24]:

$$S(t) = \frac{M_0 \sin(\theta) e^{-\mathrm{TE} R_2^*(t)} (1 - e^{-\mathrm{TR} R_1(t)})}{(1 - \cos(\theta) e^{-\mathrm{TR} R_1(t)})} \tag{10}$$

where $S(t)$ is the voxel signal intensity at time t, M_0 is the equilibrium magnetization of the protons, θ is the flip angle, TR is repetition time, and TE is echo time. $R_1(t)$ is the time course of longitudinal relaxation rate ($R_1 = 1/T_1$) and $R_2^*(t)$ is time course of transverse relaxation rate (R_2^*=$1/T_2^*$). For TE $<< T_2^*$, $e^{-\mathrm{TR} R_2^2(t)} \approx 1$.

The time course of measured signal is converted to the T_1 relaxation rate. Under the fast exchange limit, the change in R_1 ($\Delta R_1 = R_1 - R_{10}$) is linearly proportional to the tissue CA concentration C_t (mM).

$$R_1(t) = R_{10} + r_1 C_t(t) -> \Delta R_1(t) = r_1 C_t(t) \tag{11}$$

where R_{10} is the pre-contrast relaxation rate R_1.

2.6 | DCE-MRI Analysis

The $C_t(t)$ was fitted to extended Tofts model (ETM) to estimate the physiological parameters. The C_t for the ETM is given by [5]:

$$C_t(t) = \mathrm{K}^{trans} \int_0^t e^{-k_{ep}(t-\tau)} C_p(t) d\tau + C_p v_p \tag{12}$$

where $C_p(t)$ is the time course of the plasma concentration of the CA (i.e., called arterial input function [AIF]), K^{trans} (min^{-1}) is the volume transfer constant (vascular space to the EES), and $k_{ep} = \mathrm{K}^{trans}/v_e(\mathrm{min}^{-1})$ is the rate constant for CA transport from the EES to vascular space. AIF for each patient was selected from a location at the carotid artery [25]. Regions of interest (ROI's) on tumors were identified and contoured manually by an experienced neuroradiologist based on T_2w images and late phases of the T_1w DCE images. Tumor volume was calculated from the T_2w images.

Pre-contrast T1 values (ie, T_{10}) were calculated from the precontrast T1w variable flip angle images [26]. The time course of tissue concentration dataset, $C_t(t)$, was fitted to Equation (12) using a nonlinear fitting technique that minimizes the sum of squared errors (SSE) between the model fit and data. Parameter estimation values were bounded in the fitting routine as follows: $\mathrm{K}^{trans} \in [0, 5]$ (min^{-1}), v_e, and $v_p \in [0, 1]$. The ETM parametric maps were generated using in-house MRI-QAMPER software (Quantitative Analysis Multi-Parametric Evaluation Routines) written in MATLAB.

2.7 | OMT Numerical Simulation

We now outline our numerical implementation of the regularized OMT problem based on the work of [27]. We first note that, while we can numerically solve the optimization problem in a constrained manner, it is more efficient to eliminate the ADE (7) and solve the problem for v alone. Thus, we solve the PDE for a given v, and obtain the differentiable map

$$F(v) = \mu(t, x), \quad t \in [0, 1]. \tag{13}$$

Suppose the given images are $(m_1 \times \ldots \times m_d)$ in size, let $s = m_1 \times \ldots \times m_d$ denote the total number of voxels, and let K denote the number of time steps such that $K * \delta t = T$. As proposed in [27], we first employ operator splitting to discretize the ADE (7) as an advection step and diffusion step, independently. First, we consider the advection equation and solve the problem

$$\frac{\partial \mu}{\partial t} + \nabla \cdot (\mu v) = 0 \quad \mu(t_n, x) = \mu_n.$$

Using a particle in cell (PIC) method, one obtains the discrete equivalent

$$\mu^*_{n+1} = \mathbf{P}(v_n)\mu_n$$

where $\mathbf{P}$ is a linear interpolation matrix. The method is conservative, that is, no mass is lost during this step.

Even though we take the diffusion operator in the present paper to be of the form $\sigma^2\Delta$, the methodology works the same for a more general operator $\nabla \cdot D\nabla$ where D is a symmetric, positive definite matrix. According, we consider the diffusion equation and solve the problem

$$\frac{\partial \mu}{\partial t} = \nabla \cdot D\nabla\mu \quad \mu(t_n, x) = \mu^*_{n+1}.$$

Using the backward Euler method, we derive the discrete equivalent

$$(\mathbf{I} - \delta t \mathbf{A})\mu_{n+1} = \mu^*_{n+1}, \tag{14}$$

where $\mathbf{I}$ is the identity matrix and $\mathbf{A}$ is a discretization of the diffusion operator $\nabla \cdot D\nabla$ on a cell centered grid. Combining the above discretizations, we obtain the corresponding discrete forward problem $(\mathbf{I} - \delta t \mathbf{A})\mu_{n+1} = \mathbf{P}(v_n)\mu_n,\ n = 0, \ldots, K$. Note that the density at any given time step only depends upon the initial density μ_0 and the velocity v, allowing us to define the discrete map $F(v)$ (13) that maps the velocity to the density at all times. Next, setting $\mu = [\mu_1^\top, \ldots, \mu_{K+1}^\top]^\top$ and $v = [v_0^\top, \ldots, v_K^\top]^\top$, a straightforward discretization of the energy gives

$$\int_0^T \int_A \mu \|v\|^2 dx\, dt \approx h^d \delta t \boldsymbol{\rho}^\top (\mathbf{I}_K \otimes \mathbf{A}_v)(v \odot v), \tag{15}$$

where h^d is the volume of each cell, $\mathbf{I}_k$ is the $k \times k$ identity matrix, $\mathbf{A}_v$ is a $1 \times d$ block matrix of $\mathbf{I}_s$, $\otimes$ denotes the Kronecker product and $\odot$ denotes the Hadamard product. Following [27], we solve the following discrete optimization problem:

$$\min \Xi(v) = h^d\, \delta t \mu^\top (\mathbf{I}_K \otimes \mathbf{A}_v)(v \odot v) + \epsilon^2 \|\mu_{K+1} - \mu_{K+1}^{\text{obs}}\|^2$$
$$\text{subject to} \begin{cases} (\mathbf{I} - \delta t \mathbf{A})\mu_{n+1} - \mathbf{P}(v_n)\mu_n = 0 \\ \mu_0 = \mu_0^{\text{obs}}. \end{cases}$$

We note that the objective function is quadratic with respect to v and the interpolation matrix $\mathbf{P}$ is linear with respect to v since it contains the weights on the linear interpolation. Thus one can employ a Gauss-Newton like method to solve the problem. Full details may be found in [27].

Recall that the GR-OMT algorithm yields the velocity characterizing the evolution between two given images, subject to the advection-diffusion equation (7). As described above, this velocity is influenced by both advective and diffusive behavior of the CA and looking at the average speed, (i.e. the magnitude of the velocity), turns out to be quite revealing. In particular, we define *forward flux* (Φ_F, min^{-1}) to be the average speed of the velocity at some initial time interval (CA delivery phase) and *backward flux* (Φ_B, min^{-1}) to be the average speed of the velocity for the remaining time interval; the selection of the the specific time intervals for Φ_F and Φ_B is explained in the Results section. The flowchart diagram for the data processing steps in the GR-OMT pipeline is illustrated in Figure 3.

A significant advantage of OMT framework is that we also get directional information from the velocity which cannot be inferred from imaging alone. In accordance with the finding that contrast is advectively driven out of the tumor and enters the tumor via diffusion (because of differences in elevated tumor pressure [16]), a voxel is considered to be *advective* if the time-varying velocity vector at the corresponding location points out of the tumor for a majority of the time steps and *diffusive* if the velocity vector points outward for a majority of the time steps. The *diffusive percentage* is then defined to be the number of diffusive voxels divided by the tumor volume.

3 | RESULTS

For DCE-MRI images, we refer to the i-th image as the image taken at time i. We selected time intervals of $t = 6 - 21$ for forward flux and $t = 21 - 39$ for backward based on the observation that the diffusion percentage spikes at $t = 21$ across all our patients when looking at smaller time intervals/increments, as shown in Figure 2; the spike indicates the strong reflux (characteristic of backward flux). Figure 3 shows the OMT pipeline run where the corresponding forward

and backward fluxes and their respective advection/diffusion components are illustrated. The mean tumor forward flux in this case is 0.0098 and the diffusive percentage in the delivery phase ($t = 6 - 21$) is 14.21% and increases to 22.66% in the remaining interval.

Figure 6 shows comparison of ETM-derived K^{trans} and OMT-derived Φ_F. Abrupt changes in the neighboring slices/voxels are observed for K^{trans} since it ignores the intervoxel movement. In contrast, the Φ_F shows clearer signal strength and maintains the integrity of the neighboring voxels. As observed, K^{trans} follows similar trends to Φ_F.

We show forward and backward fluxes with corresponding advection and diffusion components for two patient datasets in Figures 4 and 5 based on the tumor size. Quantitative metrics for forward and backward flux for all the patients (used in this study) with and without diffusion are given in Table 1. The corresponding diffusive percentages for all the patients are shown in Table 2. Representative volumes for Figures 4 and 5 are 24.91cm^3 and 28.80cm^3, respectively.

Finally, the sensitivity analysis of different values of σ^2 is given in Figure 7. We see that the direction of the flux changes gradually as σ^2 increases but remains stable. The movement becomes more uniform as the value increases with a large difference observed when $\sigma^2 = 1$, as illustrated in (f). Given the stability of the low σ^2 values, we used $\sigma^2 = 0.0001$ for all the HNSCC datasets in this paper.

4 | DISCUSSION

In this paper, we have described a natural fluid flow model applied to DCE-MRI head and neck data. The model, unlike ETM, allows to take into account both advective and diffusive flows between neighboring voxels. In our framework, we directly applied an advection/diffusion model to study the optical flow properties of the dynamic imagery with no extraneous parameters. Thus the methodology described in the present work is data driven. We have shown how the bulk motion is chained to the diffusion through the diffused Hamilton-Jacobi equation (22). The relation to the compartmentalized models are revealed in (6) that shows explicitly how the time course of CA concentration is effected by a general velocity term that captures the conservation principles underlying the compartmental models.

Tofts-Kermode model takes as input, AIF, a transient signal change in one of the pre-selected arterial voxels and assumes these changes are the same for all the voxels in the image. We do not make any such assumption, however, AIF for Tofts-Kermode and our analysis did give the same boundary, i.e. at $t = 21$ for all the HNSCC datasets. ETM provides estimate of K^{trans} based on the non-linear fitting with initial guess values along with the lower and upper bounds, which might not provide the optimal parameter values.

There is another way of looking at our approach, which may illuminate some of the key ideas. Optimal mass transport defines a metric on distributions that allows natural physically based interpolations of data. Mathematically, it defines a geodesic path in the space of distributions in a certain precise sense [28]. However, it does not take into account diffusion. Adding diffusion to the model not only regularizes the OMT framework, but adds another important interpretation to our proposed method. Namely, (2) may be regarded as a Fokker-Planck equation, and hence describes the transition probabilities of the associated stochastic differential equation [29]. Thus, instead of moving along geodesic paths, we are now moving along Brownian paths [19]. This is a very powerful point of view in problems in which there is noise, and when one does not want to impose any *a priori* constraints. This is exactly the type of treatment underpinning data driven methodologies, which are coming to dominate many parts of applied science and engineering.

In the future, we plan to estimate the pressure directly from the presented approach. We will also add source terms to our formulation that will allow us to combine OMT with other metrics such as L^2 and Fisher-Rao [30] while still keeping the underlying advection/diffusion model governing the flow. Further, we will consider more general versions

of the diffusion coefficient, e.g. taking into account spatial dependence in order to further improve our scheme for the kinetic modeling of tumor DCE data.

ACKNOWLEDGEMENTS

This project was supported by AFOSR grant (FA9550-17-1-0435), grants from National Institutes of Health (R01-AG048769, R01-CA198121), MSK Cancer Center Support Grant/Core Grant (P30 CA008748), and a grant from Breast Cancer Research Foundation (grant BCRF-17-193).

CONFLICT OF INTEREST

There are no conflicts of interest.

REFERENCES

[1] King AD, Vlantis AC, Tsang RKY, Gary TMK, Au AKY, Chan CY, et al. Magnetic Resonance Imaging for the Detection of Nasopharyngeal Carcinoma. American Journal of Neuroradiology 2006;27(6):1288–1291. http://www.ajnr.org/content/27/6/1288.

[2] King AD, Mo FK, Yu KH, Yeung DK, Zhou H, Bhatia KS, et al. Squamous cell carcinoma of the head and neck: diffusion-weighted MR imaging for prediction and monitoring of treatment response. Eur Radiol 2010;20(9):2213–20. https://www.ncbi.nlm.nih.gov/pubmed/20309553.

[3] Lenz M, Greess H, Baum U, Dobritz M, Kersting-Sommerhoff B. Oropharynx, oral cavity, floor of the mouth: CT and MRI. Eur J Radiol 2000;33(3):203–15. https://www.ncbi.nlm.nih.gov/pubmed/10699737.

[4] Zima AJ, Wesolowski JR, Ibrahim M, Lassig AA, Lassig J, Mukherji SK. Magnetic resonance imaging of oropharyngeal cancer. Top Magn Reson Imaging 2007;18(4):237–42. https://www.ncbi.nlm.nih.gov/pubmed/17893589.

[5] Tofts PS, Kermode AG. Measurement of the blood-brain barrier permeability and leakage space using dynamic MR imaging. 1. Fundamental concepts. Magnetic resonance in medicine 1991;17(2):357–367.

[6] Shukla-Dave A, Lee NY, Jansen JFA, Thaler HT, Stambuk HE, Fury MG, et al. Dynamic Contrast-Enhanced Magnetic Resonance Imaging as a Predictor of Outcome in Head-and-Neck Squamous Cell Carcinoma Patients With Nodal Metastases. International Journal of Radiation Oncology*Biology*Physics 2012;82(5):1837 – 1844. http://www.sciencedirect.com/science/article/pii/S0360301611004597.

[7] Kim S, Loevner LA, Quon H, Kilger A, Sherman E, Weinstein G, et al. Prediction of response to chemoradiation therapy in squamous cell carcinomas of the head and neck using dynamic contrast-enhanced MR imaging. AJNR Am J Neuroradiol 2010;31(2):262–8. http://www.ncbi.nlm.nih.gov/pubmed/19797785.

[8] Jackson A, Li KL, Zhu X. Semi-quantitative parameter analysis of DCE-MRI revisited: monte-carlo simulation, clinical comparisons, and clinical validation of measurement errors in patients with type 2 neurofibromatosis. PLoS One 2014;9(3):e90300. https://www.ncbi.nlm.nih.gov/pubmed/24594707.

[9] Huang W, Chen Y, Fedorov A, Li X, Jajamovich GH, Malyarenko DI, et al. The impact of arterial input function determination variations on prostate dynamic contrast-enhanced magnetic resonance imaging pharmacokinetic modeling: a multicenter data analysis challenge. Tomography: a journal for imaging research 2016;2(1):56.

[10] Othman AE, Falkner F, Kessler DE, Martirosian P, Weiss J, Kruck S, et al. Comparison of different population-averaged arterial-input-functions in dynamic contrast-enhanced MRI of the prostate: effects on pharmacokinetic parameters and their diagnostic performance. Magnetic resonance imaging 2016;34(4):496–501.

[11] Sourbron SP, Buckley DL. On the scope and interpretation of the Tofts models for DCE-MRI. Magnetic resonance in medicine 2011;66(3):735–745.

[12] Sourbron S, Buckley DL. Tracer kinetic modelling in MRI: estimating perfusion and capillary permeability. Physics in Medicine & Biology 2011;57(2):R1.

[13] Rachev S, Rüschendorf L. Mass Transportation Problems. Springer-Verlag, New York; 1998.

[14] Villani C. Topics in Optimal Transportation. No. 58, American Mathematical Soc.; 2003.

[15] Benamou JD, Brenier Y. A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem. Numerische Mathematik 2000;84:375–393.

[16] Ewing JR, Nagaraja TN, Aryal MP, Keenan KA, Elmghirbi R, Bagher-Ebadian H, et al. Peritumoral tissue compression is predictive of exudate flux in a rat model of cerebral tumor: an MRI study in an embedded tumor. NMR in Biomedicine 2015;28(11):1557–1569.

[17] Elkin R, Nadeem S, Haber E, Steklova K, Lee H, Benveniste HB, et al. GlymphVIS: Visualizing Glymphatic Transport Pathways Using Regularized Optimal Transport. MICCAI 2018 2018;.

[18] Yankeelov TE, Lepage M, Chakravarthy A, Broome EE, Niermann KJ, Kelley MC, et al. Integration of quantitative DCE-MRI and ADC mapping to monitor treatment response in human breast cancer: initial results. Magnetic resonance imaging 2007;25(1):1–13.

[19] Chen Y, Georgiou TT, Pavon M. On the relation between optimal transport and Schrödinger bridges: A stochastic control viewpoint. Journal of Optimization Theory and Applications 2016;169(2):671–691.

[20] Peyré G, Cuturi M. Computational Optimal Tranport. ArXiv:1803.00567, 2018; 2018.

[21] Li W, Yin P, Osher S. Computations of optimal transport distance with Fisher information regularization. Journal of Optimization Theory and Applications 2018;75(3):1581–1595.

[22] Liero M, Mielke A, Savare G. Optimal transport in competition with reaction: the Hellinger–Kantorovich distance and geodesic curves. SIAM Journal Math Analysis 2016;48:2869–2911.

[23] Paudyal R, Oh JH, Riaz N, Venigalla P, Li J, Hatzoglou V, et al. Intravoxel incoherent motion diffusion-weighted MRI during chemoradiation therapy to characterize and monitor treatment response in human papillomavirus head and neck squamous cell carcinoma. Journal of Magnetic Resonance Imaging 2017;45(4):1013–1023.

[24] Bagher-Ebadian H, Jain R, Nejad-Davarani SP, Mikkelsen T, Lu M, Jiang Q, et al. Model selection for DCE-T1 studies in glioblastoma. Magn Reson Med 2012;68(1):241–51. `https://www.ncbi.nlm.nih.gov/pubmed/22127934`.

[25] Shukla-Dave A, Lee N, Stambuk H, Wang Y, Huang W, Thaler HT, et al. Average arterial input function for quantitative dynamic contrast enhanced magnetic resonance imaging of neck nodal metastases. BMC Med Phys 2009;9:4. `http://www.ncbi.nlm.nih.gov/pubmed/19351382`.

[26] Deoni SC, Peters TM, Rutt BK. High-resolution T1 and T2 mapping of the brain in a clinically acceptable time with DESPOT1 and DESPOT2. Magn Reson Med 2005;53(1):237–41. `https://www.ncbi.nlm.nih.gov/pubmed/15690526`.

[27] Steklova K, Haber E. Joint hydrogeophysical inversion: state estimation for seawater intrusion models in 3D. Computational Geosciences 2017;21(1):75–94.

[28] Otto F. The geometry of dissipative evolution equations: the porous medium equation. Communications in Partial Differential Equations 2001;26.

[29] Gardiner C. Stochastic Methods: A Handbook for the Natural and Social Sciences. Springer; 2008.

[30] Chen Y, Georgiou T, Tannenbaum A. Interpolation of matrices and matrix-Valued measures: The Unbalanced Case. European Journal of Applied Mathematics 2018;https://doi.org/10.1017/S0956792518000219.

[31] Woodall RT, Barnes S, Hormuth DA*et al* . The Effects of Intravoxel Contrast Agent Diffusion on the analysis of DCE-MRI Data in realistic tissue domains. Magnetic Resonance in Medicine 2018;80(1):330–340.

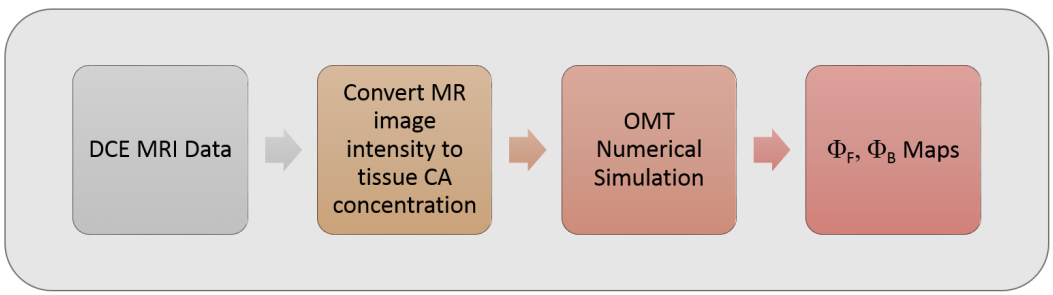
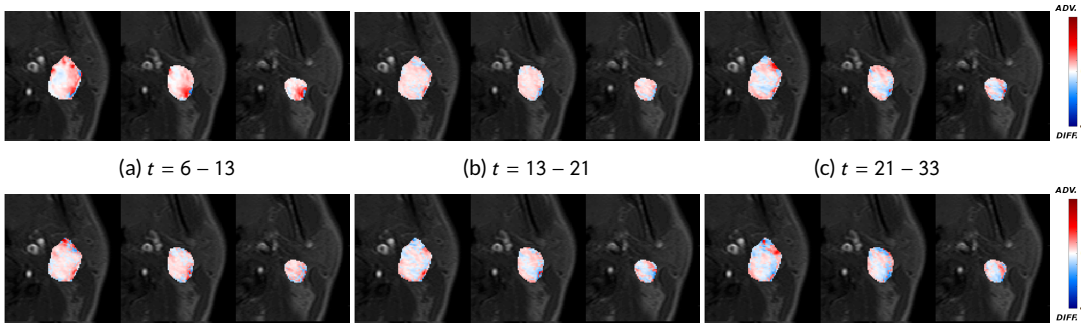


FIGURE 1 Flowchart diagram of data processing steps for OMT pipeline. DCE MRI data is acquired and image signal intensity is converted to time-course contrast agent (CA) tissue concentration data using multi-flip angle T_1w images. The CA data is interpreted as mass density in the image via the OMT simulation, which produces estimated maps of forward flux (Φ_F) and backward flux (Φ_B).



(a) $t = 6 - 13$ (b) $t = 13 - 21$ (c) $t = 21 - 33$

(a) $t = 33 - 42$ (b) $t = 42 - 51$ (c) $t = 51 - 60$

FIGURE 2 Smaller time increments. The diffusion percentages are (a) 18.76%, (b) 20.28%, (c) 25.14%, (d) 22.10%, (e) 23.66% and (f) 23.49%. We use the diffusion spike at (c) to divide the total time interval into $t = 6 - 21$ for forward flux and $t = 21 - 60$ for backward flux.

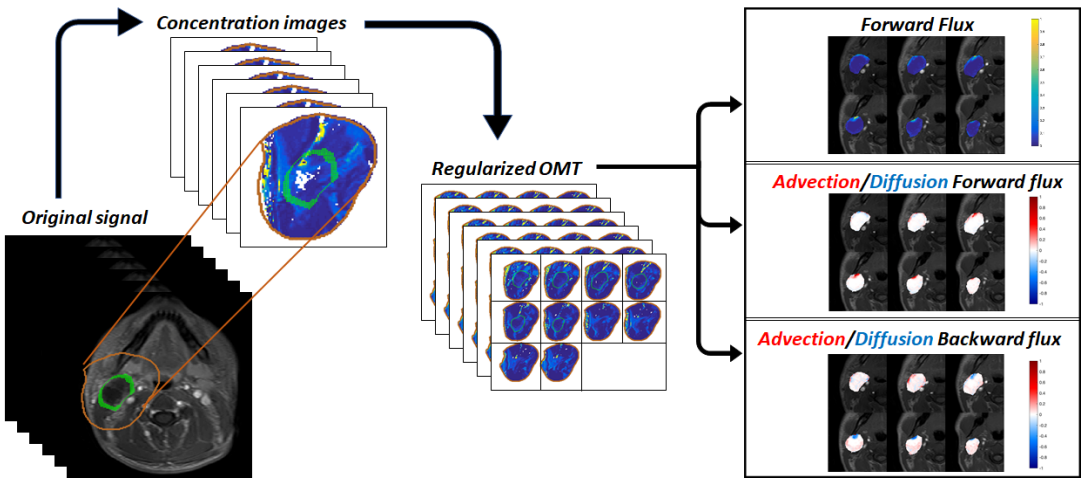


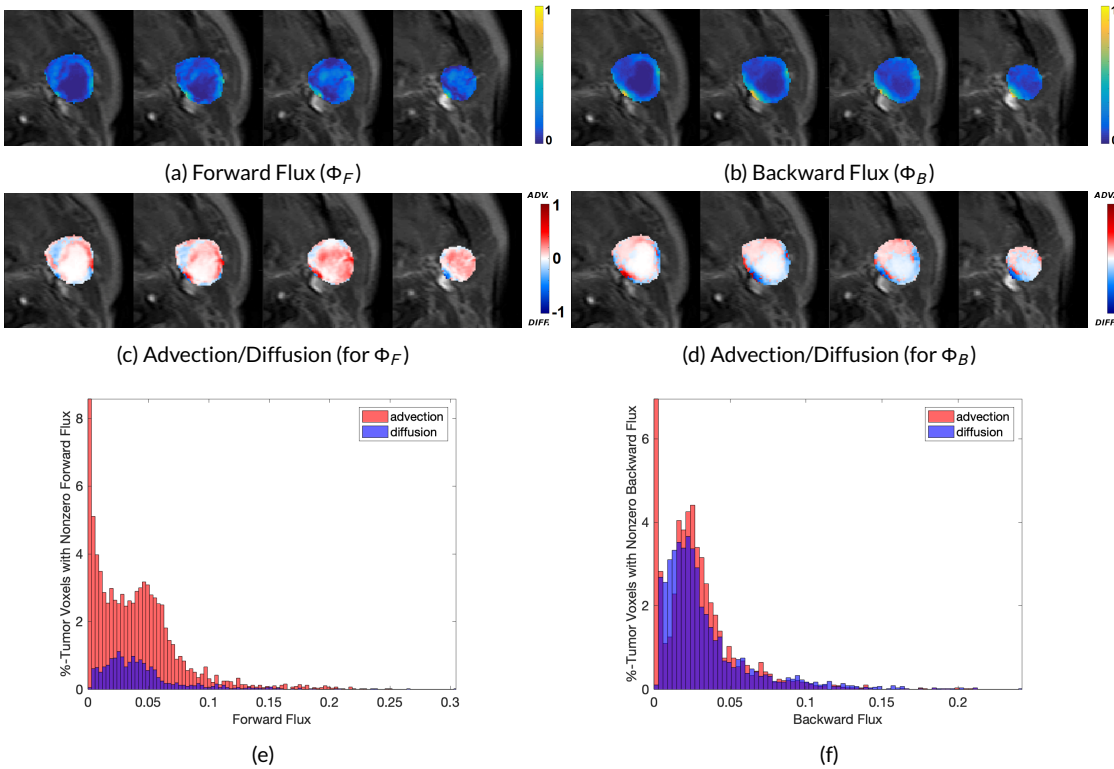
FIGURE 3 Optimal Transport pipeline. MR intensity images are converted to concentration images which are then fed to our regularized OMT algorithm. The resultant interpolations give 'clean' density maps and the corresponding velocity field is used to compute the forward flux by averaging the speed components over the delivery time points and pressure-driven advection/diffusion flows on the boundary via the directional component (advective flows are represented by movement out of the tumor whereas the diffusion is represented by movement into the tumor, i.e. against the pressure gradient).

TABLE 1 mean w/o diffusion = (sum of advective speeds)/(number of advective + diffusive voxels)

	Forward Flux (Φ_F min^{-1})				Backward Flux (Φ_B min^{-1})			
Patient ID	with diffusion		without diffusion		with diffusion		without diffusion	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
1	0.0148	0.0146	0.0099	0.0139	0.0219	0.0117	0.0102	0.0134
2	0.0405	0.0354	0.0330	0.0357	0.0311	0.0279	0.0157	0.0242
3	0.0077	0.0134	0.0063	0.0118	0.0089	0.0120	0.0061	0.0113
4	0.0221	0.0267	0.0165	0.0263	0.0176	0.0288	0.0114	0.0268
5	0.0215	0.0323	0.0151	0.0298	0.0234	0.0230	0.0131	0.0214
6	0.0470	0.0412	0.0372	0.0452	0.0438	0.0410	0.0293	0.0437
7	0.0243	0.0225	0.0167	0.0223	0.0177	0.0244	0.0089	0.0215
8	0.0082	0.0091	0.0052	0.0086	0.0103	0.0108	0.0033	0.0071
9	0.0098	0.0184	0.0068	0.0182	0.0032	0.0040	0.0014	0.0026
10	0.0107	0.0151	0.0079	0.0146	0.0124	0.0151	0.0074	0.0146

TABLE 2 Diffusive percentages in forward and backward flux for all the patients.

Patient ID	Forward Flux (Φ_F)	Backward Flux (Φ_B)
1	18.76%	27.61%
2	8.67%	17.42%
3	9.77%	14.70%
4	9.96%	20.16%
5	11.38%	15.63%
6	11.83%	16.60%
7	14.21%	23.11%
8	15.20%	30.36%
9	14.21%	22.66%
10	10.78%	12.76%



(a) Forward Flux (Φ_F)

(b) Backward Flux (Φ_B)

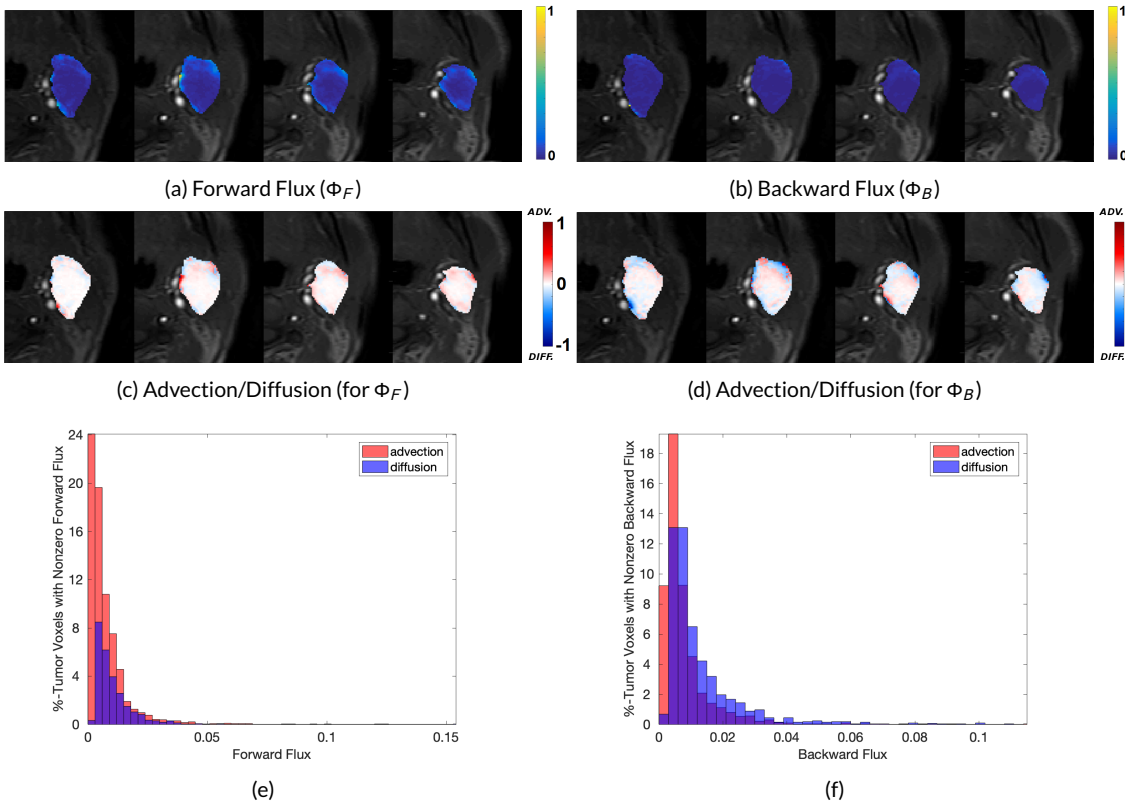
(c) Advection/Diffusion (for Φ_F)

(d) Advection/Diffusion (for Φ_B)

(e)

(f)

FIGURE 4 Forward and backward flux and the corresponding advection and diffusion components. The diffusion percentage increases from 8.67% in forward flux to 17.42% in backward flux. **(e)** Mean advective (red) and diffusive (blue) speeds given as the percentage of tumor voxels with nonzero mean speed (Forward Flux).



(a) Forward Flux (Φ_F)

(b) Backward Flux (Φ_B)

(c) Advection/Diffusion (for Φ_F)

(d) Advection/Diffusion (for Φ_B)

(e)

(f)

FIGURE 5 Forward and backward flux and the corresponding advection and diffusion components. The diffusion percentage increases from 15.20% in forward flux to 30.36% in backward flux. **(e)** Mean advective (red) and diffusive (blue) speeds given as the percentage of tumor voxels with nonzero mean speed (Forward Flux).

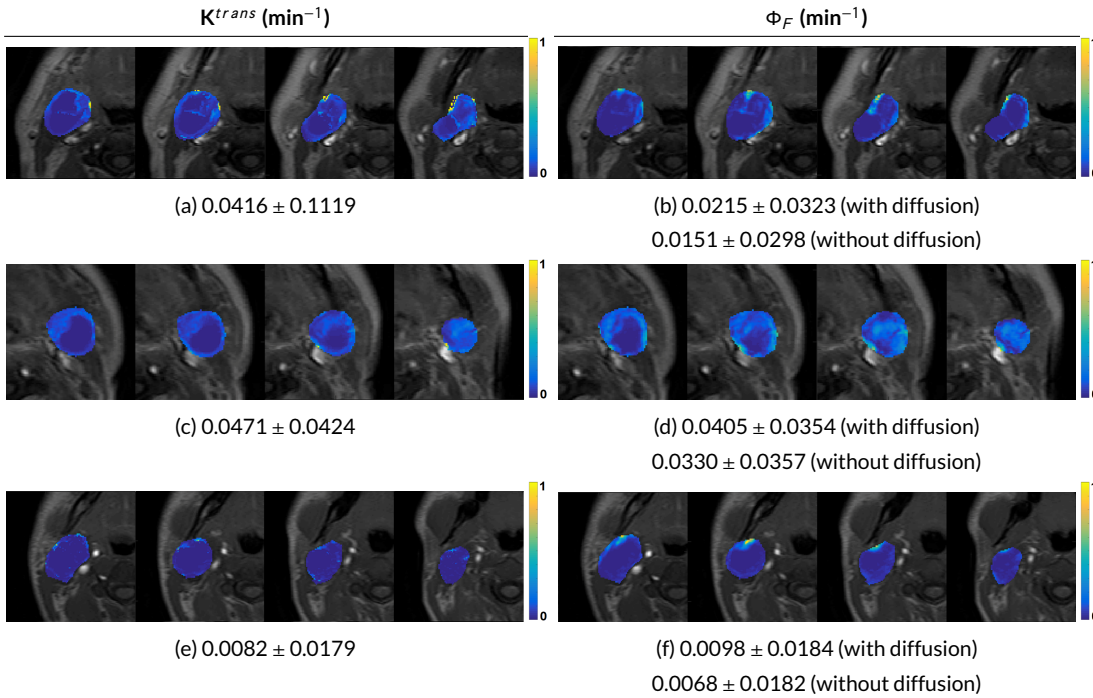
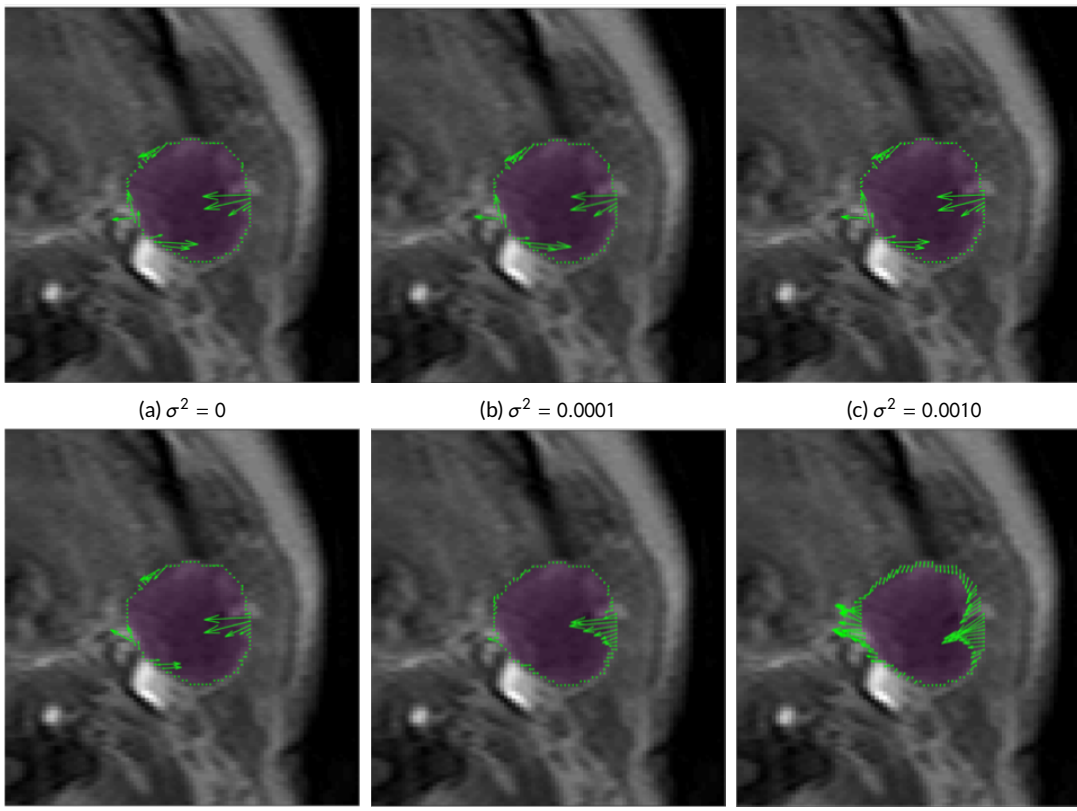


FIGURE 6 Comparison of mean±standard deviation for K^{trans} and forward flux (Φ_F) values on 4 neighboring slices for 3 HNSCC patients. Note the abrupt changes in K^{trans} neighboring slices/voxels since it ignores the intervoxel movement. In contrast, forward flux shows clearer signals and maintains the integrity of the neighboring voxels.



(a) $\sigma^2 = 0$ (b) $\sigma^2 = 0.0001$ (c) $\sigma^2 = 0.0010$

(d) $\sigma^2 = 0.0100$ (e) $\sigma^2 = 0.1000$ (f) $\sigma^2 = 1.0000$

FIGURE 7 Comparison of flux $j = \sigma^2 \nabla\mu - \mu v$ for increasing values of σ^2. The flux vectors are shown in green on the tumor boundary, highlighted in purple. We see that the direction of the flux changes gradually as σ^2 increases but remains stable. The movement becomes more uniform as the value increases with a large difference observed when $\sigma^2 = 1$, as illustrated in (f). Given the stability of the low σ^2 values, we used $\sigma^2 = 0.0001$ for all the HNSCC datasets in this paper.

SUPPLEMENTAL THEORY

Background on Optimal Mass Transport

The main tool we will be employing is based on the theory of optimal mass transport (OMT). We will not give all the details here, but simply refer the interested reader to [13, 14]. For our purposes, we will review the fluid dynamical version of OMT over the normalized time interval $t \in [0, 1]$ [15]:

$$\inf_{\mu,v} \quad \int_0^1 \int_A \frac{1}{2}\mu(t,x)\|v(t,x)\|^2 \, dx \, dt \tag{16}$$

$$\text{subject to :} \quad \partial_t \mu + \nabla \cdot (\mu v) = 0 \tag{17}$$

$$\mu(0,\cdot) = \mu_0(\cdot), \quad \mu(1,\cdot) = \mu_1(\cdot), \tag{18}$$

where $\mu = \mu(t,x) \in \mathbb{R}$ is the density interpolant between the given densities μ_0 and μ_1, assumed to have the same total mass, $v = v(t,x) \in \mathbb{R}^d$ is the velocity, and $A \subset \mathbb{R}^d$ is a subdomain. Further, we assume that $\mu_i : A \to \mathbb{R}^+$ has compact support $i = 0, 1$. For our purposes we take $d = 2$ or $d = 3$. Here, we consider the objective function given as the kinetic energy associated with the transport (16) and require that the time-varying densities and velocities satisfy the continuity equation (17) with temporal endpoints matching the given densities (18) to enforce mass conservation. It has been shown that the optimal $\mu_{\min}$ and $v_{\min}$ such that the infimum (16) is achieved do exist [15]. In simple terms, this approach seeks a flow over some time-space interval that minimizes the mass weighted kinetic energy of the flow.

Adding diffusion

Following [17], we modify the original Benamou and Brenier OMT formulation by adding a diffusion term in the continuity equation (17) to allow for ***both advection and diffusion*** of the contrast agent and having the added benefit of regularizing the flow. Various aspects of regularized OMT have studied in a number of recent works including [19, 20, 21, 22] which have extensive lists of references.

Accordingly, we formulate the following ***regularized OMT*** problem

$$\inf_{\mu,v} \int_0^1 \int_A \quad \frac{1}{2}\mu(t,x)\|v(t,x)\|^2 \, dx \, dt \tag{19}$$

$$\text{subject to} \quad \partial_t \mu + \nabla \cdot (\mu v) = \nabla \cdot (D \nabla \mu), \tag{20}$$

$$\mu(0,\cdot) = \mu_0(\cdot), \quad \mu(1,\cdot) = \mu_1(\cdot), \tag{21}$$

where D, encapsulating spatially-varying diffusion, is a positive definite symmetric matrix modelling the diffusion.

Hamiton-Jacobi equation

Using calculus of variations [14], one can show the optimal velocity for (19) is of the form

$$v_{\min} = -\nabla \lambda,$$

where λ satisfies the following diffused version of a Hamilton-Jacobi equation:

$$\partial_t \lambda - \frac{1}{2}\|\nabla\lambda\|^2 = \nabla \cdot D\nabla\lambda. \tag{22}$$

The latter equation derived from OMT (minimizing kinetic energy of the flow subject to an advection-diffusion constraint) is very revealing regarding the proposed model. Indeed, we note that the optimal velocity is the gradient of a function, and via (22), the bulk flow (advection) is clearly influenced as well by the diffusion. This will be described further in Section below.

| Tofts-Kermode Model analogy to OMT

We very briefly sketch the relation to the very popular Tofts-Kermode model [5]. For simplicity, we indicate details in 2D. The basic idea is to integrate the advection-diffusion equation (20) over a given compartment V and use the divergence theorem. Accordingly, we let S denote the boundary of V. We first note that

$$\int_V \nabla \cdot D\nabla\mu dV = \int_S D\nabla\mu \cdot \vec{\mathbf{n}} dS =: \mathcal{D}_\mu, \tag{23}$$

where $\vec{\mathbf{n}}$ denotes the outward normal to the surface S.

Remark: In [31], $\nabla\mu \cdot \vec{\mathbf{n}}$ is interpreted as the difference between the concentrations of the contrast agent in the plasma and EES weighted by the ***permeability***, defined as K^{trans} [5], divided by the total surface area of the vasculature within the voxel multiplied by the volume of tissue within the voxel. This instantiation of Ficke's serves to relate diffusive flux to differences in concentration.

Next we note that

$$\int_V \nabla \cdot (\mu v)\, dV = \int_S \mu v \cdot \vec{\mathbf{n}}\, dS \tag{24}$$

$$= -\int_S \mu\nabla\lambda \cdot \vec{\mathbf{n}}\, dS =: \mathcal{P}_\mu. \tag{25}$$

This part, derived from advection, may be interpreted as arising from Darcy's law (difference in pressure) but because of (22), there is also the influence of diffusion. We should emphasize that OMT theory gives a unique solution of (20) with optimal velocity given as the gradient of a function via (22).
Finally, setting

$$\int_V \mu dV =: C(t), \tag{26}$$

we see that at the optimal solution of the OMT problem, (20) becomes an ordinary differential equation of the form

$$\frac{dC(t)}{dt} + \mathcal{P}_\mu = \mathcal{D}_\mu. \tag{27}$$

Equation (27) leads to pharmacokinetic modelling. The compartmental integrated diffusive part is $\mathcal{D}_\mu$ (which depends on Ktrans [31] and has been utilized for its estimation), and the compartmentalized advection part (driven by bulk flow, pressure coupled with diffusion via (22)) is given by $\mathcal{P}_\mu$. Combining $\mathcal{P}_\mu$ and $\mathcal{D}_\mu$, equation (27) shows that the

time rate of change in concentration over the compartment can be written as

$$\frac{dC(t)}{dt} = \int_S (D\nabla\mu - \mu v) \cdot \vec{\mathbf{n}}\, dS, \tag{28}$$

where $\vec{\mathbf{j}} = D\nabla\mu - \mu v$ is the total flux. Although the direction of the flux, $\vec{\mathbf{j}}$, is influenced by both the density gradient $\nabla\mu$ and the velocity v, we chose to focus our attention on the velocity since it is affected by time-varying changes in the given images as well as diffusion, as seen from the Hamilton-Jacobi equation (22). Interestingly, we find that the velocity alone yields information that aligns well with the current understanding of the contrast behavior which we will elaborate on below.

| Full OMT model for flow of contrast agent

In this section, we summarize our overall model based on the regularized OMT model. In what follows, we will take the diffusion matrix $D = \sigma^2 I$, that is, a constant times the identity. In other words, diffusion is not a tensor quantity, and there are no preferred directions. The sensitivity analysis of different values of σ^2 is given in Figure 7.

We make the following assumptions about the data to motivate the proposed regularized model:

(i) Image intensity is proportional to contrast agent density. Intensity is taken as the mass density in the OMT model.

(ii) The flow is governed by the advection-diffusion equation (ADE),

$$\partial_t \mu + \nabla \cdot (\mu v) = \sigma^2 \Delta\mu, \tag{29}$$

where $0 \leq \mu : [0, T] \times A \to \mathbb{R}$ is a density with compact support, $A \subset \mathbb{R}^d$ is a subdomain, $v : [0, T] \times A \to \mathbb{R}^d$ is velocity and $0 \leq \sigma \in \mathbb{R}$ is diffusivity. As alluded to above, we take $d = 3$ in what follows. Further, in our treatment above we took $T = 1$. Here we take a general final time T. The mathematical treatment is identical.

(iii) We are given noisy observations of the contrast agent density at times $t = T_i$,

$$\mu(T_i, x) + \gamma = \mu_{T_1}^{\text{obs}}(x), \quad i = 0, \ldots, N, \tag{30}$$

where γ is Gaussian iid noise with covariance Σ.

Given initial and final observations of contrast agent density μ_0^{obs} and μ_T^{obs} at times $t = 0$ and $t = T$ respectively, we need to determine the velocity field v and the 'true' or 'clean' images μ_0 and μ_T such that the constraint (29) is satisfied. Accordingly, we propose to minimize (with respect to v and μ) the objective function

$$\int_0^T \int_A \mu \|v\|^2 \, dx dt + \epsilon^2 \|\mu(T, x) - \mu_T^{\text{obs}}(x)\|_\Sigma^2, \tag{31}$$

subject to the ADE constraint (29). The notation indicates that the variance operation is inversely scaled by the covariance terms.

Here, the second term is added to the Benamou and Brenier energy functional [15] to prevent over-fitting to noise in the data. The parameter ϵ^2 drives the balance between fitting the data and minimizing the energy associated with transporting the mass. We refer to (31) (with constraint (29)) as the ***generalized regularized OMT*** problem (GR-OMT).