Efficient DCE-MRI Parameter and Uncertainty Estimation Using a Neural Network

Yannick Bliesener, *Student Member, IEEE*, Jay Acharya, and Krishna S. Nayak, *Senior Member, IEEE*

***Abstract*— Quantitative DCE-MRI provides voxel-wise estimates of tracer-kinetic parameters that are valuable in the assessment of health and disease. These maps suffer from many known sources of variability. This variability is expensive to compute using current methods, and is typically not reported. Here, we demonstrate a novel approach for simultaneous estimation of tracer-kinetic parameters and their uncertainty due to intrinsic characteristics of the tracer-kinetic model, with very low computation time. We train and use a neural network to estimate the approximate joint posterior distribution of tracer-kinetic parameters. Uncertainties are estimated for each voxel and are specific to the patient, exam, and lesion. We demonstrate the methods' ability to produce accurate tracer-kinetic maps. We compare predicted parameter ranges with uncertainties introduced by noise and by differences in post-processing in a digital reference object. The predicted parameter ranges correlate well with tracer-kinetic parameter ranges observed across different noise realizations and regression algorithms. We also demonstrate the value of this approach to differentiate significant from insignificant changes in brain tumor pharmacokinetics over time. This is achieved by enforcing consistency in resolving model singularities in the applied tracer-kinetic model.**

***Index Terms*— Quantitative imaging, DCE MRI, parameter estimation, uncertainty estimation.**

I. Introduction

DYNAMIC contrast enhanced magnetic resonance imaging (DCE-MRI) aims to estimate sub-voxel parameters of pathology through fitting of tracer-kinetic (TK) models to contrast agent concentration-time-curves [1]. As with other quantitative MRI techniques, clinical applicability of DCE-MRI measurements hinges on accurate, precise, and reproducible results [2]–[4]. Many of the involved cost functions from non-linear tracer-kinetic models are, however, not strictly convex or not even convex. This introduces ambiguity in the parameters since many parameter settings fit the measured data equally well, and further deteriorates accuracy and precision due to susceptibility to noise and initialization [5]–[8]. Development of such estimators is furthermore challenged by lack of reference methods and ground truth to appropriately assess these crucial figures of merit.

One solution is non-Bayesian maximum-likelihood estimation (MLE). Such methods typically use optimization algorithms and multiple initializations to solve issues arising from regions of complete or approximate flatness in the cost function landscape [6], [9], [10]. More recently, neural networks have been proposed to shift the computational load from (frequent) inference to (one-time) training [11]. For these algorithms, measures of precision are obtained through either time-consuming Monte-Carlo simulations (MC) with multiple noise realizations and initializations, or variance estimation through linear error propagation [12].

Another solution is Bayesian estimation, which reduces ambiguity and variance in the estimated tracer-kinetic parameters by introducing prior information [5]. Measures of precision can simply be derived from the posterior distribution [6], [8]. The application of such estimators has two important limitations. First, one must specify a suitable prior distribution on the parameters to be estimated. While a good prior distribution can significantly improve the estimation quality, a poor choice might deteriorate the result as well [5]. Second, these algorithms are based on Bayes' rule, which involves finding the partition function. Numerous approaches exist in the literature and have been applied to DCE-MRI. One approach is to obtain the maximum-a-posteriori estimate (MAP) by optimization methods without computation of the partition function, which further requires Monte-Carlo simulations to study bias and variance [5], [6]. A second approach is to linearize the observation model and to approximate the posterior by a Gaussian distribution which enables analytic solutions of the iterative update steps in the Expectation-Maximization algorithm; yet successful convergence may require proper initialization and restarts [8], [13], [14]. In other approaches, access to the posterior distribution is obtained through numerical integration of the partition function [6], or by sampling from it through Markov-Chain Monte-Carlo methods [10], [15]–[18]. Alternatively, the integration can be simplified by choice of conjugate-priors [19]. All of these methods either involve substantial computation time [8], [15]–[17], suffer from convergence issues [13], or limit the scope of posteriors [8], [13], [14] or priors [19] that can be used.

Manuscript received September 30, 2019; accepted November 8, 2019. Date of publication November 26, 2019; date of current version April 30, 2020. Research reported in this publication was supported by the National Cancer Institute of the National Institutes of Health under award number R33-CA225400. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. *(Corresponding author: Yannick Bliesener.)*

Y. Bliesener and K. S. Nayak are with the Ming Hsieh Department of Electrical and Computer Engineering, University of Southern California, Los Angeles, CA 90089 USA (e-mail: bliesene@usc.edu; knayak@usc.edu).

J. Acharya is with the Department of Radiology, Keck School of Medicine, University of Southern California, Los Angeles, CA 90033 USA (e-mail: Jay.Acharya@med.usc.edu).

Color versions of one or more of the figures in this article are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TMI.2019.2953901

In this work, we propose the use of a neural network to estimate approximate posterior distribution of tracer-kinetic parameters. From these posterior distributions, ranges of tracer-kinetic parameters are derived that allow one to estimate the parameters as well as make decisions about confidence in the estimation. While being computationally fast, the proposed approach is flexible in the choices of priors and applicable posterior assumptions. As a prior distribution we chose the non-informative uniform prior. This naturally renders the estimation task an application of Jaynes' maximum entropy principle [20], which maximizes the entropy of the estimated posterior distribution subject to known constraints, e.g., data-consistency in this case. We then demonstrate application of parameter ranges instead of point measurements in treatment response assessment during longitudinal monitoring of brain tumor patients at our institution.

II. Tracer-Kinetic Modeling

DCE-MRI monitors wash-in, accumulation, and wash-out of a T_1-shortening contrast agent [1]. The measured four-dimensional (three spatial dimensions and one time dimension) signals are converted into concentration-time-curves $C(t)$ (CTC) for each voxel inside the field of view. This is done by use of previously measured baseline maps for longitudinal relaxation T_{10} and proton density M_0 as outlined in [21]. While this conversion is in principle a non-linear, complex mapping [22], typical parameter settings, i.e., high imaging flip angle, short echo time, and low dose of bolus, allow for a linear approximation on the real numbers which maintains the additive-Gaussian nature of the noise, yet entails different noise power in each voxel [23], [24].

In parametric DCE-MRI, the voxel concentration-time-curves are fit to compartmental models to estimate parameters of local physiology. While many models exist [1], this paper focuses on the extended Tofts-Key model (ETK) which is one of the most widely used models for brain tumors [25], [26]. For this model the concentration $C(t)$ is governed by

$$C(t) = \Phi\left(\theta; C_\mathrm{p}\right) = \mathrm{v_p}\, C_\mathrm{p}(t) + \mathrm{K^t} \int_0^t C_\mathrm{p}(\tau)\, \mathrm{e}^{-\frac{\mathrm{K^t}}{\mathrm{v_e}}(t-\tau)}\, \mathrm{d}\tau. \quad (1)$$

Here, the arterial input function $C_\mathrm{p}(t)$ (AIF) models the bolus delivery to the compartments over time t, $\mathrm{v_p}$ and $\mathrm{v_e}$ are relative volumes of blood plasma and extra-vascular-extracellular space (EES), respectively, and $\mathrm{K^t}$ models the transfer of contrast agent from plasma to EES [1].

III. Bayesian Estimation

A. Variational Approach to Posterior Approximation

The starting point for Bayesian estimation of a parameter vector $\theta = \left(\mathrm{K^t}, \mathrm{v_p}, \mathrm{v_e}\right)$ is the posterior distribution $p(\theta|C)$ which is given by Bayes' Rule:

$$p(\theta|C) = \frac{p(C|\theta)\, p(\theta)}{\int p(C|\zeta)\, p(\zeta) d\zeta}, \quad (2)$$

where $p(C|\theta)$ denotes the data model and $p(\theta)$ is the prior distribution on θ. As discussed in Section II, the data model is given by a normal distribution

$$p(C|\theta) = \mathcal{N}\left(\Phi\left(\theta; C_\mathrm{p}\right), \sigma^2\right). \quad (3)$$

Similar to previous work on Bayesian estimation for DCE-MRI we deploy independent uniform distributions for all three tracer-kinetic parameters of the extended Tofts-Kety model [6], [17]. Hence, the model does not assume any physiologically induced correlations between different tracer-kinetic parameters. Examples of such correlations are tumors which are known to have proliferated vasculature ($\mathrm{v_p}$) while simultaneously exhibiting increased leakiness ($\mathrm{K^t}$) like pleomorphic xanthoastrocytomas [27]. The posterior is, however, able to capture correlations in θ due to intrinsic collinearity of the tracer-kinetic model.

To avoid integration in (2), we choose a variational approach that approximates the posterior by minimizing the Kullback-Leibler divergence between the approximating distribution $q(\theta|C)$ and the posterior $p(\theta|C)$ [19], [28]:

$$\min_\Psi \mathcal{L} = \min_\Psi D\left(q_\Psi(\theta|C)||p(\theta|C)\right) \quad (4)$$

$$= \min_\Psi \mathbb{E}_q\left[-\log\left(\frac{p(C|\theta)\, p(\theta)}{q_\Psi(\theta|C)\, p(C)}\right)\right], \quad (5)$$

where Ψ denotes the parameter vector for the approximate posterior distribution function, e.g., the borders of support for a uniform distribution. Neglecting constant terms and applying data model and the prior for the tracer-kinetic parameters the cost function becomes

$$\min_\Psi \frac{1}{2\sigma^2}\mathbb{E}_q\left[\|C - \Phi(\theta; C_\mathrm{p})\|_2^2\right] + \mathbb{E}_q\left[\log\left(q_\Psi(\theta|C)\right)\right], \quad (6)$$

where C and C_p are measured concentration-time-curve and arterial input function, respectively. σ is the (measured) noise standard deviation.

The first term in (6) enforces data consistency of the tracer-kinetic parameters from the approximated posterior distribution. The second term in (6) resembles a negative differential entropy and enforces spread of the posterior distribution to encompass all parameters that could potentially explain the data. By implicitly maximizing entropy while enforcing a data consistency constraint this cost formulation is a special case of Jaynes' Maximum Entropy Principle [20]. Hence, the combination of variational approximation and (uninformative) uniform prior on the tracer-kinetic parameters provides a clear strategy how to compute the posterior distribution in case of ambiguity. Such ambiguity may result from unidentifiability of tracer-kinetic parameters in the low contrast concentration regime and under excessive noise corruption.

Contour plots of the data consistency cost function exhibit valleys around the optimum with approximate elliptical shape (cf. Fig. 5 and Figs. 4a to 4c). A bias analysis [5] further shows bias for $\mathrm{K^t}$ and $\mathrm{v_e}$ in (1) for many realistic parameter combinations (see Section SI of the supplementary material). The approximation of the posterior distribution should therefore be capable of predicting parameters inside ellipsoidal geometries without given preference to the lowest value for the data consistency as this may lead to false suggestion during inference. With regard to these observations,

we chose the posterior distribution to be uniform within ellipsoidal boundaries which further facilitates to avoid non-physiological negative parameter ranges and can easily adapt to non-symmetric valleys in the cost function landscape. This choice however adds a further layer of approximation to the method as the solution space of Jaynes' Maximum Entropy Problem is known to be the exponential family of distributions which does not contain the uniform distribution [29].

During the optimization process of (6) the expectation operation for the data consistency term is replaced by an empirical average. Random samples Θ of $q_\Psi(\theta|C)$ are drawn according to the following rule:

$$\Theta = \mathrm{R}(\Psi)\mathrm{S}(\Psi)\mathbf{X} + \mathbf{m}(\Psi), \tag{7}$$

where $\mathbf{X} \in \mathbb{R}^3$ are uniform samples from the unit ball B. $\mathrm{R}(\Psi) \in \mathbb{R}^{3\times3}$ is a rotation matrix, $\mathrm{S}(\Psi) \in \mathbb{R}^{3\times3}$ is a diagonal scaling matrix, and $\mathbf{m}(\Psi) \in \mathbb{R}^3$ is a vector adjusting the center of the ellipsoid.

Since differential entropy is translation and rotation invariant, the second term in (6) can easily be computed from the differential entropy $H(B)$ of the uniform unit ball as [30]

$$\mathbb{E}_q\left[\log\left(q_\Psi(\theta|C)\right)\right] = -H(B) - \log\det \mathrm{S}(\Psi). \tag{8}$$

B. Estimation of Posterior Parameters

One of the objectives of this work is to make Bayesian estimation computationally feasible in clinical settings. Neural networks present a way to explicitly express functionality that would otherwise be only implicitly available such as in form of solutions to an optimization problem (6). This allows neural networks to shift the main computational load away from frequent inference to training, which is typically only done once.

Intrinsic to parametric DCE-MRI is the coding of many time frames describing the local dynamics of the contrast agent into few static parameter values, e.g., $\mathrm{K^t}$, $\mathrm{v_p}$, and $\mathrm{v_e}$ in the case of maximum likelihood estimation, or the parameters Ψ of the posterior distribution of $\mathrm{K^t}$, $\mathrm{v_p}$, and $\mathrm{v_e}$ in the present work. A similar task is achieved by the encoder part of an autoencoder when used for dimensionality reduction or denoising [31], [32]. For this reason, we adapt the encoder structure consisting of fully connected layers with decreasing numbers of connections as main component of the neural network architecture [11], [31], [32].

To guide encoding, Ulas *et al.* [11] deployed a multiscale (spatial) filter bank which feeds the features it recognizes into the encoder structure. Following this approach, we preconnect to the encoder two identical banks of temporal filters, one for the arterial input function input and one for the concentration-time-curve input, each with three different lengths and strides: 1) A filter with maximum length and stride. Inspired by computation of cerebral blood volume, the ratio of the output of this filter for concentration-time-curve and arterial input function would allow to compute $\mathrm{v_p}$ in the absence of contrast leakage [1]. 2) Measurements of backflux of contrast agent are typically observed on longer time scales [14], [33]. This motivates the use of filters with broader temporal scope and stride to aid estimation of $\mathrm{v_e}$. 3) Finally, a filter with short length and stride is added to allow capture of fast dynamics at the early phase after bolus arrival. Accurate capture of the initial phase of the arterial input function is essential for tracer-kinetic parameter estimation [34]. To be able to handle different onset times identically we fix the stride of this filter to 1.

The lengths and strides of the second and third filter as well as number of encoder layers and the respective widths were tuned to the task at hand. A description of the tuning procedure and plots to illustrate the tuning strategy can be found in Section SII of the supplementary material.

All but the last dense layer are LeakyReLU-activated, while the last layer is activated with a sigmoid. The sigmoid activation normalizes the output of the network to lie in the unit-cube, thus facilitating re-parameterization to respective meaningful parameter ranges. The complete two-stage neural network architecture is illustrated in Fig. 1.

The uniform-ellipsoid posterior distribution can be parameterized in multiple ways. For the 3-parameter extended Tofts-Kety model (1), the parameterization of the posterior $\Psi \in \mathbb{R}^9$ could naturally be chosen as: three translation, three scaling, and three rotation parameters. The first data consistency term in (6) affects all of Ψ, the negative entropy loss (8), however only the scaling entries in Ψ. To aid training of the network we therefore parameterize the ellipsoids by the six extreme points of their major axes and three rotation angles. To further lower over-parameterization introduced by rotational symmetry we confine the rotation angles to lie in $\left[-\frac{\pi}{4}, \frac{\pi}{4}\right]$.

With regard to the cost function in (6), the network performance is naturally tied to the specific level of noise σ it has been trained for. Because the conversion of signal-time-curves to concentration-time-curves entails a scaling of each voxel by a factor that depends on the baseline T_{10} and M_0 values of that voxel, the noise level for each of the concentration-time-curves will be different. We therefore scale arterial input function and concentration-time-curves on a per voxel basis to match the noise level of the network prior to the first filter stage.

C. Training Data

We synthesized training data based on the extended Tofts-Kety model (1). Seventy-one arterial input functions were manually selected from clinical exams at our institution and scaled to fall into the range $[0, 1]$ mmol/L. With these arterial input functions, 511,200 concentration-time-curves where generated by randomly drawing tuples of $(\mathrm{K^t}, \mathrm{v_p}, \mathrm{v_e})$. Two third of the voxels were in the range $[0, 0.6]\,\mathrm{min}^{-1} \times [0, 0.6] \times [0.05, 0.6]$ representing brain tumor physiology [26], [27], [35]. The remaining one third was $\mathrm{K^t} = 0$, $\mathrm{v_e} = 0.5$, and $\mathrm{v_p}$ drawn at random from $[0, 0.6]$ to mimic healthy tissue. We generated a total of 1,533,600 concentration-time-curves by using three noise realizations per patch with $\sigma = 0.001$. Arterial input functions were randomly shifted with shifts in the interval of $[-4, +4]$ samples and scaled by a factor in the interval $[0.01, 1]$ before generation of concentration-time-curves. This trains the network to be invariant to onset times and to handle a range of SNR levels. Note that arterial input functions are synchronized with the concentration-time-

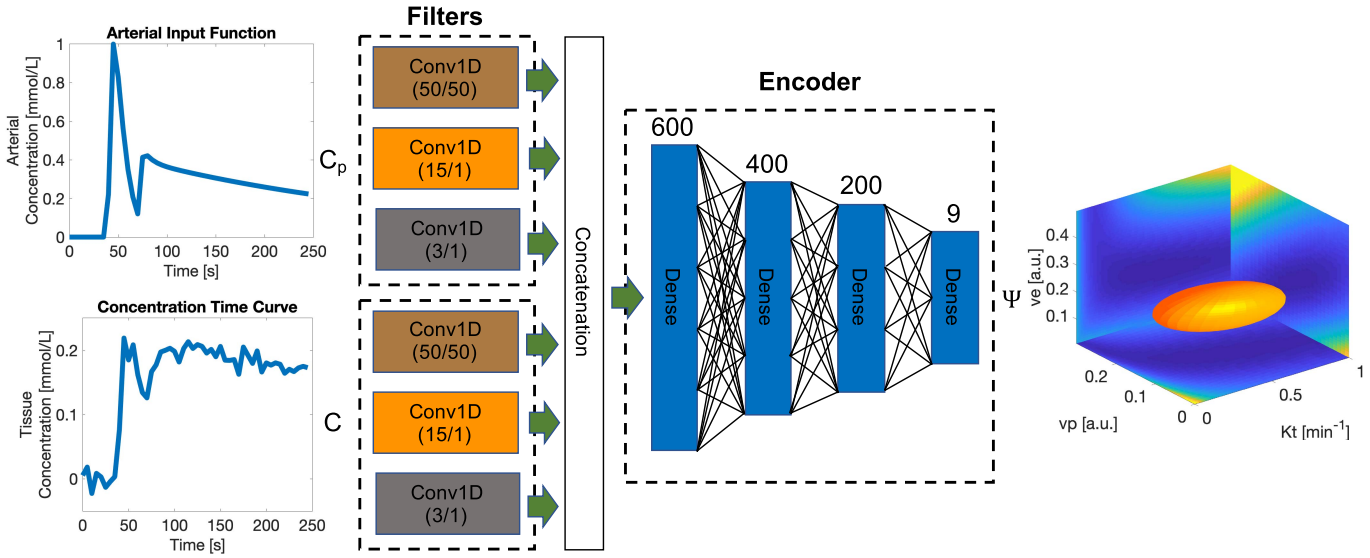

Fig. 1. The neural network consists of two stages: An input filter stage for $C_p(t)$ and $C(t)$, and an encoder stage. The input filter stage comprises three temporal filters for each $C_p(t)$ and $C(t)$. Numbers in parentheses show filter length and stride. After concatenation the outputs of the filter stage are encoded into 9 parameters for the posterior distribution through a sequence of 4 dense layers. Numbers count neurons of the dense layers. Every layer in the network except the last is activated with LeakyReLU. The last has a sigmoid activation. The output of the last layer are 9 parameters for scale, rotation, and translation of the approximate joint posterior of K^t, v_p, and v_e.

curves for each voxel prior to the estimation, i.e., zero onset delay, but enhancement onset may be different for each concentration-time-curve.

D. Regularization and Training Procedure

At low contrast leakage, i.e., low values of K^t or low values of v_e, the data consistency loss tends to form two valleys in an L-shaped form of which only one gives physiologically realistic values for K^t and v_e (an example of this can be seen in Fig. 4a). To force the network to inflate the distribution ellipsoid into the desired trench, we make use of the known true labels for the training data and extend the cost function in (6) to force the bounding box of the ellipsoid to encompass the true parameter. Together with (8) this yields:

$$\min_{\Psi} \frac{1}{2\sigma^2}\mathbb{E}_q\left[\|C-\Phi(\theta;C_p)\|_2^2\right]-\log\det S(\Psi) \\ +\lambda\, d(\Psi,\theta_{\text{true}}), \quad (9)$$

where

$$d(\Psi,\theta_{\text{true}}) = \inf_{x\in\boxed{\Psi}} \|x-\theta_{\text{true}}\|_2 \quad (10)$$

is the distance of the true parameter θ_{true} to the inscribed box $\boxed{\Psi}$ of the ellipsoid parameterized by Ψ.

We further regularize the the convolutional filters in the input filter stage of the neural network to have unit l_2-norm and by total-variation penalty. Additionally, dropout layers were added after every layer of the network in Fig. 1 and the weights of the dense layers are regularized with l_1-norm penalty. We enforce axis-alignment of the posterior ellipsoid in voxels mimicking healthy tissue because this improved estimation accuracy in non-enhancing tissue at the lower end of the SNR range.

To achieve fast training, we designed a 3-stage procedure: First, the network is trained for 400 epochs with Adam optimizer with standard settings [36] and exponential learning rate decay. During this stage each element in the batch is weighted by the inverse of its noise variance before averaging to yield the cost function value. Second, the training is restarted with Adam optimizer with standard settings and exponential rate decay for 400 epochs, yet the elements of the batch are weighted by the inverse of the noise standard deviation. Last, the cost function weighting is removed and training continues for 600 epochs.

IV. Simulation Study

A. Digital Reference Object

To establish a ground truth for parameter maps, we use an anatomically and physiologically realistic digital reference objects [37], which consists of a patch of glioblastoma tumor (Fig. 2d and 2g). The parameter maps for the digital reference object were generated using the identical pipeline outlined in [37]. This entails fitting of the extended Tofts-Kety model to clinical data and bi-linear interpolation of voxels that were detected as fitting failures. Concentration-time-curves are generated for tumor patches and 50 time frames with 5s temporal resolution following the extended Tofts-Kety model (1) with an arterial input function that is extracted from one of our clinical scans that is not part of the training data. When arterial input functions and concentration-time-curves are scaled such that the arterial input functions fall in the range $[0,1]$ mmol/L, noise measurements across all datasets used in this study showed 2$^{\text{nd}}$ and 98$^{\text{th}}$ percentile of noise levels in tumor concentration-time-curves to be at $\sigma_{2\text{prctile}} = 0.002$ and $\sigma_{98\text{prctile}} = 0.05$, respectively. Hence, we further add

three different levels of additive-white-Gaussian noise with $\sigma = 0.002, 0.01, 0.05$ to the concentration-time-curves.

B. Experiments

We repeat two experiments on the digital reference object for each noise level. First, we compare the true parameter maps to the mean of predicted posterior distribution by the proposed approach as well as results of a regression with a Gauss-Newton algorithm with 21 initializations for $k_{ep} = K^t/v_e$ equally spaced in the interval $[0, 0.1]\,\mathrm{min}^{-1}$. We further compare the standard deviations and observed K^t ranges of this Monte-Carlo simulation with the predicted standard deviations and K^t ranges in the tumor ROI. While the statistical meaning of these two approaches to quantify uncertainty differs, previous work on tracer-kinetic modeling has shown the observed range during Monte-Carlo simulations to be approximately equal to the support of the posterior [18].

Second, we compare the predicted standard deviations and parameter ranges in the tumor ROI with the standard deviations and parameter ranges across different regression algorithms, i.e., Levenberg-Marquardt algorithm [9], l-BFGS [21], Gauss-Newton [38], [39], Nelder-Mead Simplex method [40], and initializations which is known to be one of the sources of error preventing reproducibility and comparability of DCE MRI results [9], [41], [42]. We first specify 21 initializations for $k_{ep} = K^t/v_e$ equally spaced in the interval $[0, 0.1]\,\mathrm{min}^{-1}$. To generate different numbers of initializations, we then downsample this vector yielding 21, 10, 4, 2, and 1 initialization. For this experiment the noise realization is fixed. Results of these regressions are shown in the supplementary material in Section SIII.

C. Results

Fig. 2 summarizes spatial maps of K^t for the different fitting algorithms and different noise levels, as well as concentration-time-curves for the voxels of highest K^t. Corresponding error metrics are listed in Table. I. The proposed method demonstrates lower maximum error in estimating K^t across the full range of noise levels. The maximum K^t error of the proposed method (red dot in the spatial maps in Fig. 2) is always observed in the enhancing tumor tissue regions, and for different true K^t values at different noise levels. In terms of MSE, the proposed method is able to outperform Gauss-Newton regression at low to medium noise levels, yet exhibits increased MSE at the upper end of the noise range.

Fig. 3 shows correlation plots of predicted standard deviations in the tumor ROI with the standard deviation of the reference method over 100 noise realizations, and the standard deviation across different fitting routines. The same comparison is performed for predicted and observed parameter ranges, i.e., the range from highest to lowest estimate of K^t in each voxel.

In the low noise regime, the proposed method tends to underestimate uncertainty in K^t due to noise and differences in regression algorithms. At medium noise level corresponding to the range from 20^{th} to 80^{th} percentile of noise levels, the proposed method is able to accurately predict variability induced by different regression algorithms, while it overestimates the noise standard deviation. In the strong noise regime, the proposed method overestimates the noise induced uncertainty especially at voxels with high variability. With regards to variability induced by regression algorithm there is no observable trend.

TABLE I
ROOT MEAN SQUARE ERROR (RMSE) AND MAXIMUM ERROR OF K^T IN DIGITAL REFERENCE OBJECT EXPERIMENT FOR DIFFERENT LEVELS OF NOISE σ

	RMSE $[\mathrm{min}^{-1}]$		Max Error $[\mathrm{min}^{-1}]$	
σ	Reference	Proposed	Reference	Proposed
0.002	0.008	0.007	0.31	0.05
0.005	0.021	0.013	0.64	0.05
0.010	0.024	0.020	0.38	0.11
0.050	0.055	0.068	0.29	0.23

Fig. 4 shows contour plots of the cost function for voxels with highest K^t error in the predicted parameter map. The highest error always happens at low enhancement, i.e., at low values of K^t, when the cost function becomes non-convex. The proposed method is able to predict a parameter range that contains the true parameter, yet this range tends to be overly conservative at medium to high noise.

V. LONGITUDINAL BRAIN DCE MRI STUDY

A. Data Acquisition

We illustrate application of the proposed method to the longitudinal monitoring of brain tumors in humans. Retrospective data were obtained under a protocol approved by our Institutional Review Board. All datasets were acquired according to the clinical Brain MRI with Contrast protocol at our institution, which includes DCE-MRI, and is routinely performed in brain tumor patients. Data were anonymized prior to analysis. Informed Consent was not required for this study. For T_1 mapping Variable-Flip-Angles scans with flip angles FA $= 2°, 5°, 10°, 20°, 25°$ were acquired and used [43]. The DCE scan consisted of 50 time frames of 6 axial slices with 5s temporal resolution acquired with two-fold SENSE acceleration. Each slice has matrix size 186×256 with voxel dimension 1mm $\times$ 1mm $\times$ 7mm and was acquired with FA $= 15°$ and TE/TR $=$ 2ms/6.3ms.

B. Tumor Cases in Dataset

The dataset encompasses three lesions in three subjects: One meningioma with no treatment and no change over two time points 382 days apart, a pituitary adenoma with radiation treatment and no change over two time points 371 days apart, and a metastasized melanoma treated with gamma knife stereotactic radiosurgery, which increased in size over three exams with intervals of 130 and 151 days and regressed at the fourth 143 days later, based on standard bi-directional assessment [44].

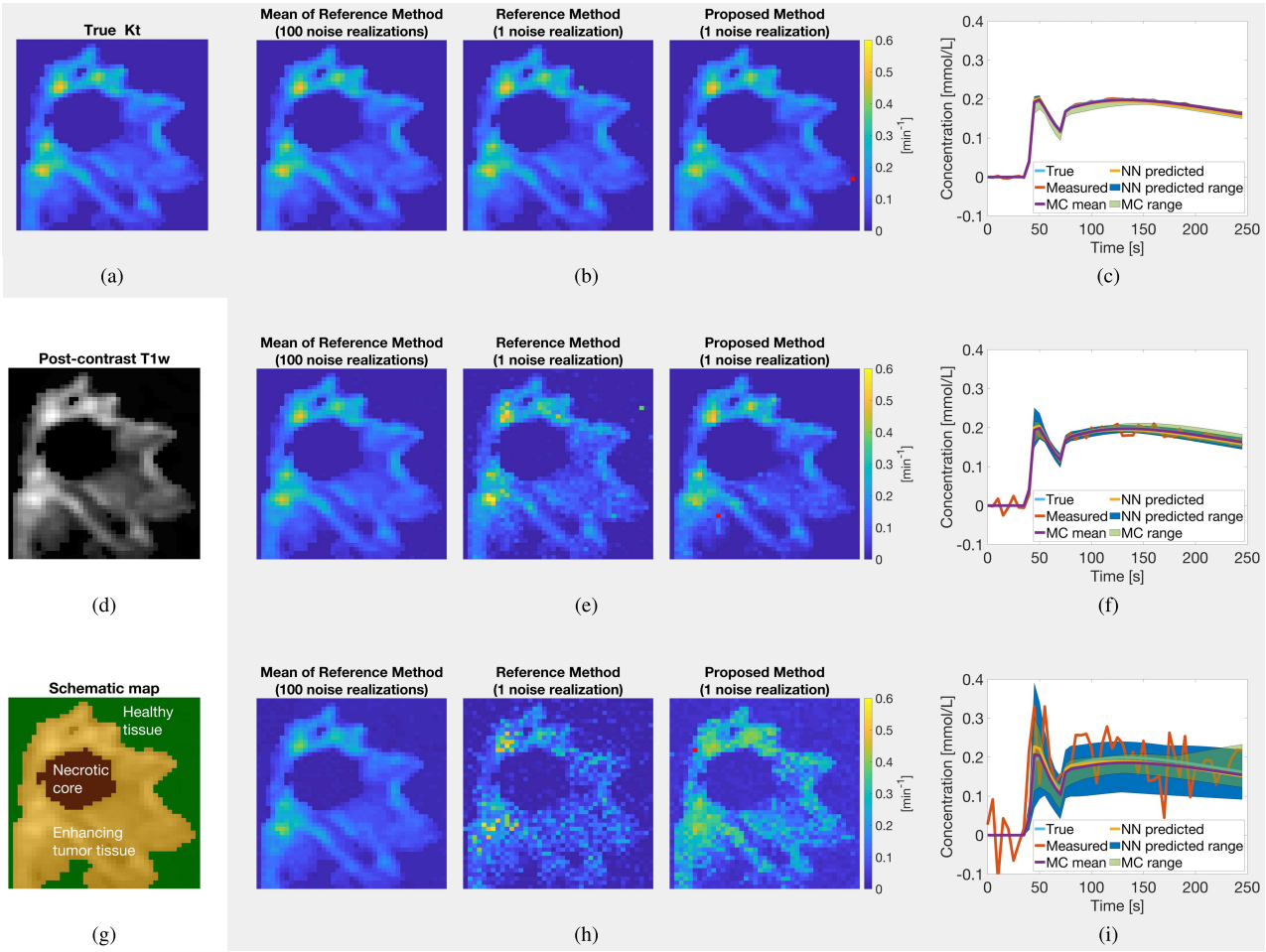

Fig. 2. Comparison of true parameter maps of the digital reference object with reference method and proposed neural network approach (NN). Each row corresponds to a different noise level $\sigma = 0.002, 0.01, 0.05$. Panel (b), (e), and (h) show the mean of reference method for Monte-Carlo simulations (MC) with 100 noise realizations, and reference method and mean of posterior distribution predicted by proposed approach for one noise realization. Parameter maps are in good agreement for low to medium noise levels. Position of highest K^t error for the proposed method is marked by a red dot. Right most column: Concentration-time-curves (CTC) for the voxel with highest K^t. While the proposed parameter map differs substantially from true and mean MC map at $\sigma = 0.05$, the predicted concentration-time-curve range and Monte-Carlo concentration-time-curverange are well within the noise floor for the most enhancing voxel.

C. Arterial Input Function Extraction

Arterial Input Functions were extracted in a semi-automatic manner. We first normalize the signal-intensity-curves to the interval $[0, 1]$. We then use the following characteristic features of arterial input functions to generate vessel masks: difference of peak enhancement to final enhancement and difference of peak to baseline signal larger than 0.2, time-to-peak and full-width-at-80%-maximum less than half the entire exam duration, respectively [45]–[47]. After erosion of each vessel in the mask to reduce partial volume averaging at the boundary of the vessel, we manually select one vessel per patient such that the signal enhancements are approximately consistent in scale across multiple exams. We assume a constant vessel $R_{10} = 0.63\text{s}^{-1}$ [48]. To account for scanner gain calibration, we estimate proton density for the vessel through a fit to baseline, and use the 95^{th} percentile of the proton density for all voxels in the vessel. After conversion to concentration-time-curves using constant values for T_1 relaxivity $r_1 = 4.5\text{s}^{-1}$ and Hematocrit Hct $= 0.42$ [49], the potential arterial input functions with peak value within top 80% of maximum are selected and averaged.

D. Reference Methods

All DCE exams were processed with conventional regression which consists of Gauss-Newton fitting of the extended Tofts-Kety model in (1) [38], [39]. These regressions are initialized with a previous Patlak fit for v_p and K^t, i.e., a regression of (1) with $v_e \to \infty$ [1], and 21 initializations for $k_{ep} = K^t/v_e$ equally spaced in $[0, 0.1]\,\text{min}^{-1}$.

To establish a diagnostic reference, all tumor cases were assessed by bi-directional measurement of the lesion on axial T1-weighted post-contrast images on a PACS workstation by a neuroradiologist with eight years of experience [44].

E. Bolus Arrival Time Estimation

Bolus arrival time was estimated voxelwise as the time point of steepest slope in a local linear approximation of

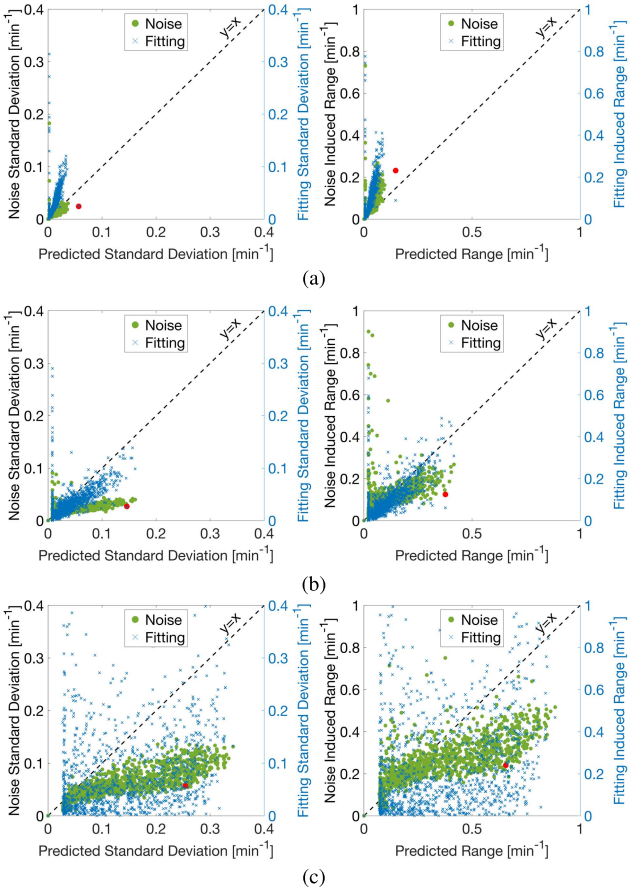

Fig. 3. Comparison of predicted and observed uncertainty of K^t. Left column: Predicted standard deviations of K^t by proposed method and observed standard deviations across different noise realizations and fitting methods. Right column: Predicted parameter ranges of K^t by proposed method and observed parameter ranges induced by noise and differences in fitting methods. Each row (a)–(c) corresponds to a different noise level $\sigma = 0.002, 0.01, 0.05$. Red dot marks voxel with highest K^t error of proposed method.

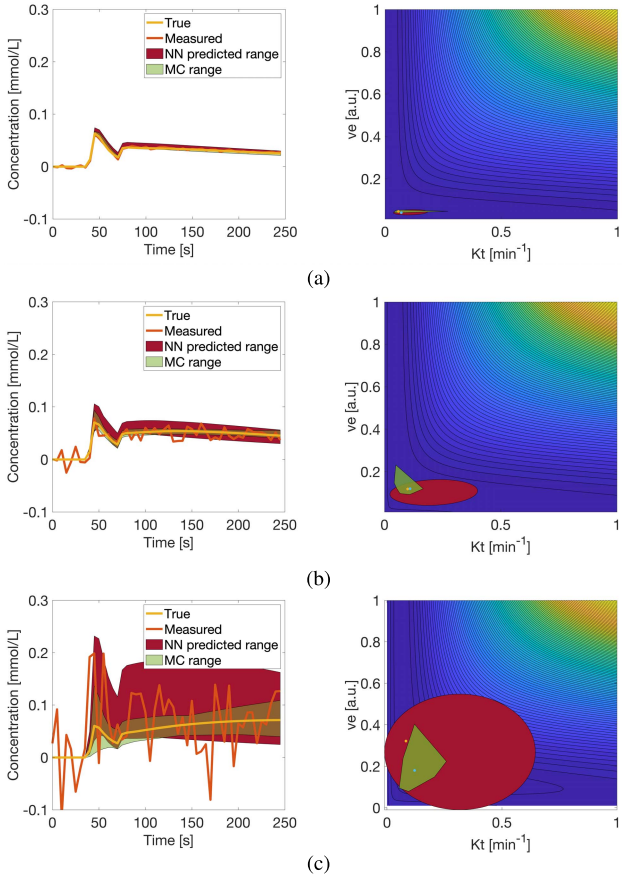

Fig. 4. Analysis of error cases for different noise levels. Left column shows true and noisy concentration-time-curves for the voxel with highest K^t error of the proposed method. These are compared with the range of concentrations observed during Monte-Carlo simulations (MC) and the predicted range by the proposed method. Right column shows the corresponding contour plots of the cost function for the K^t-v_e plane. The yellow dot indicates true parameters, blue dot the true minimum of cost function, the red ellipse the predicted posterior distribution, and the convex hull of estimated parameters during Monte-Carlo simulations is drawn in green. Each row (a)–(c) corresponds to a different noise level $\sigma = 0.002, 0.01, 0.05$.

the concentration-time-curve [50]. This bolus arrival time was used to synchronize arterial input function and concentration-time-curve in each voxel. To account for bolus arrival time estimation inaccuracies and jitter, we further performed the fitting three times with no and plus/minus one time shift of the arterial input function about the estimated bolus arrival time and report the best fit in terms of RMSE. This was done for the reference method as well as for the proposed method.

F. Biomarkers for Treatment Response

Various metrics have been used as biomarkers in DCE-MRI to assess response to treatment [51]. Many of those are based on statistics that are derived from histograms of K^t which has been shown by tumor growth models to be the most informative parameter in tumor physiology [52]. Since higher order statistics led to inconsistent trends regarding their relationship to treatment response [53], [54], yet there appears to be indication that K^t values decrease in case of positive treatment response [55], we adapt the mean K^t in the tumor ROI as biomarker [15].

G. Data Analysis

1) Verification of Approximate Posterior Distributions: Fig. 5 shows a comparison of measured concentration-time-curves and concentration-time-curves predicted by the reference method and the proposed approach as well as contour plots of sample cost function landscapes along the three hyper-planes in the parameter space. Fig. 5a shows that discrepancies between proposed and reference method can arise in the low SNR regime due to different treatment of cost function non-convexity. In Fig. 5b, the posterior ellipsoid expands into the valleys of the cost function and leads to a range of concentration-time-curves that cover the measured concentration-time-curve tightly.

2) Longitudinal Assessment of Therapy Response: Fig. 6 illustrates an exemplary post-contrast magnitude image with tumor ROI and vessel ROI for arterial input function extraction for the meningioma case at the first visit, and compares K^t maps generated with the reference method and the proposed approach. The reference method is initialized multiple times,

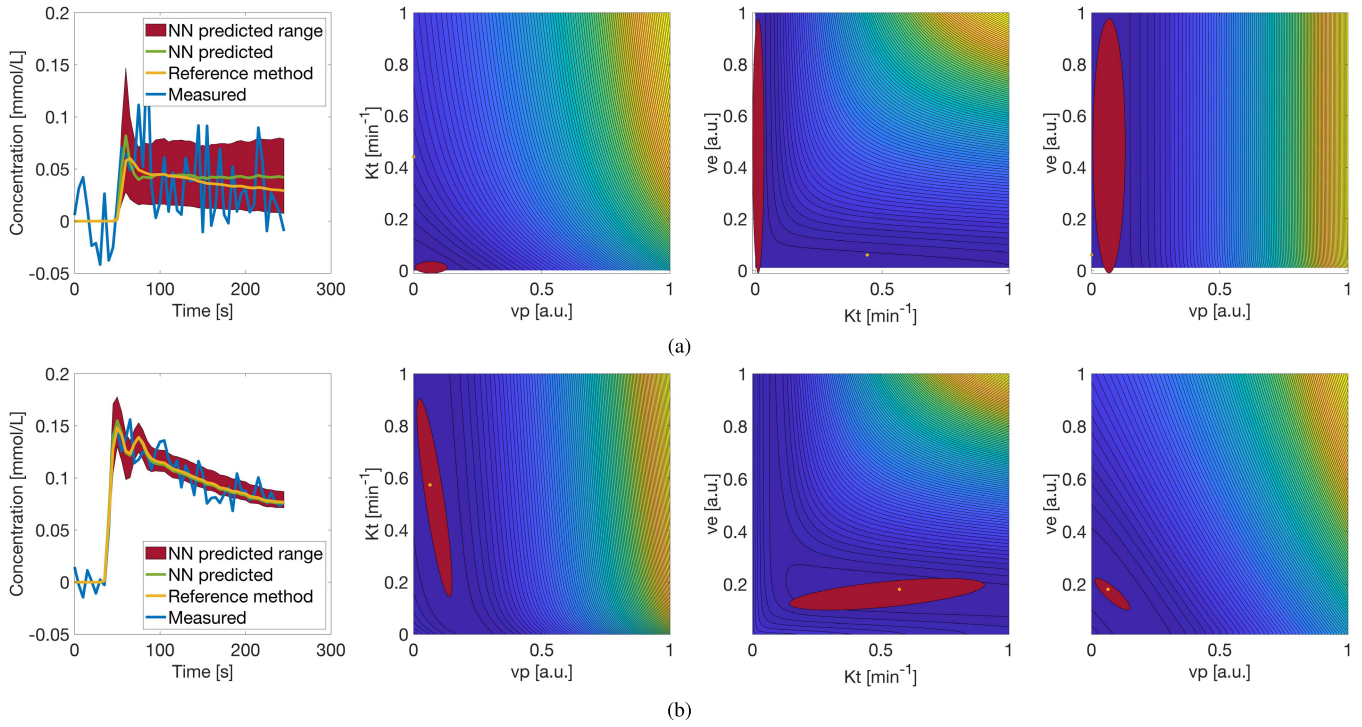

Fig. 5. Comparison of measured and fitted concentration time curves (CTC), and contour plots of respective cost function in all three hyperplanes of the parameter space. Top row (a) corresponds to the red circle in Fig. 7a, bottom row (b) to Fig. 7d. Blue curves show the measured voxel CTC. Red areas show cross sections of the neural network (NN) predicted posterior approximation which leads to CTCs inside the corresponding red swath. The mean of the posterior ellipsoid gives rise to the green CTC. Yellow dot and curve indicate the reference value and corresponding CTC. In the first case, the network predicts low enhancement due to low permeability K^t, while the reference method estimates low enhancement due to small v_e. In the second case the posterior ellipsoid is expanded into the cost function valley leading to predicting a large range of possible K^t values.

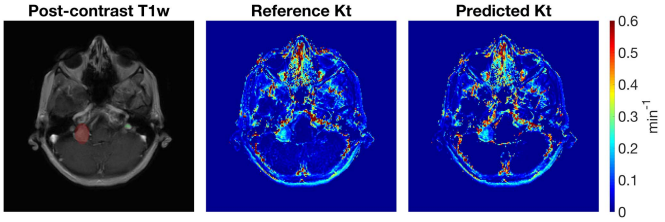

Fig. 6. Comparison of estimated K^t maps for different methods. Left-most panel shows tumor ROI (red) and arterial input function vessel (green) in post-contrast magnitude image. Middle and right-most panel compare K^t maps for reference method with multiple initialization and K^t maps estimated with the proposed method.

while the mean of the predicted parameter range is shown for the proposed approach. Overall the parameter maps are in good agreement at locations that are known to follow the extended Tofts-Kety model (1), while discrepancies in the map are visible around the skull and nasal cavities.

Fig. 7 shows predicted ranges of K^t for the first exams for all three lesions in the respective tumor ROIs, and temporal evolution of predicted ranges of K^t and average K^t in the tumor ROIs with upper and lower bounds predicted by the proposed approach.

Predicted parameter ranges are in good agreement with the reference methods for all three lesions. For Lesion 1&3 the reference method with multiple initializations shows larger values of K^t for some voxels which stem from differences in treatment of cost function non-convexity (Fig. 5a).

Temporal evolutions of K^t ranges in tumor ROIs (Figs. 7b, 7e and 7h) are in excellent agreement with the reference diagnoses.

For Lesion 1 and 3 the temporal evolution of the mean K^t in the tumor ROI matches the diagnosed tumor development. For Lesion 2, the difference in mean K^t would indicate a progression of the tumor, yet endowing the measurement with predicted ranges of uncertainty indicates the increase likely to be due to measurement errors. Many cases of pituitary adenoma (Lesion 2) can be a difficult type of tumor to clinically measure and evaluate over time. Observing the confidence intervals for all three lesions overlap substantially suggests the mean K^t lacks in specificity as biomarker to detect change in tumors.

3) Computation Time: Training time for the network was approximately 20h on a Dual NVIDIA K40. The inference of 6 slices of clinical data took approximately 10s on an office laptop without GPU acceleration. This inference was performed three times per dataset to correct for errors in the bolus arrival estimation.

VI. Discussion

We present a method to predict tracer-kinetic parameter maps and estimate their uncertainty with low computation time. Uncertainties are estimated for each voxel and are specific to patient, exam, and lesion. This is in strong contrast to other lines of work in which global bounds on the uncertainty of the measurement procedure are determined once prior to the

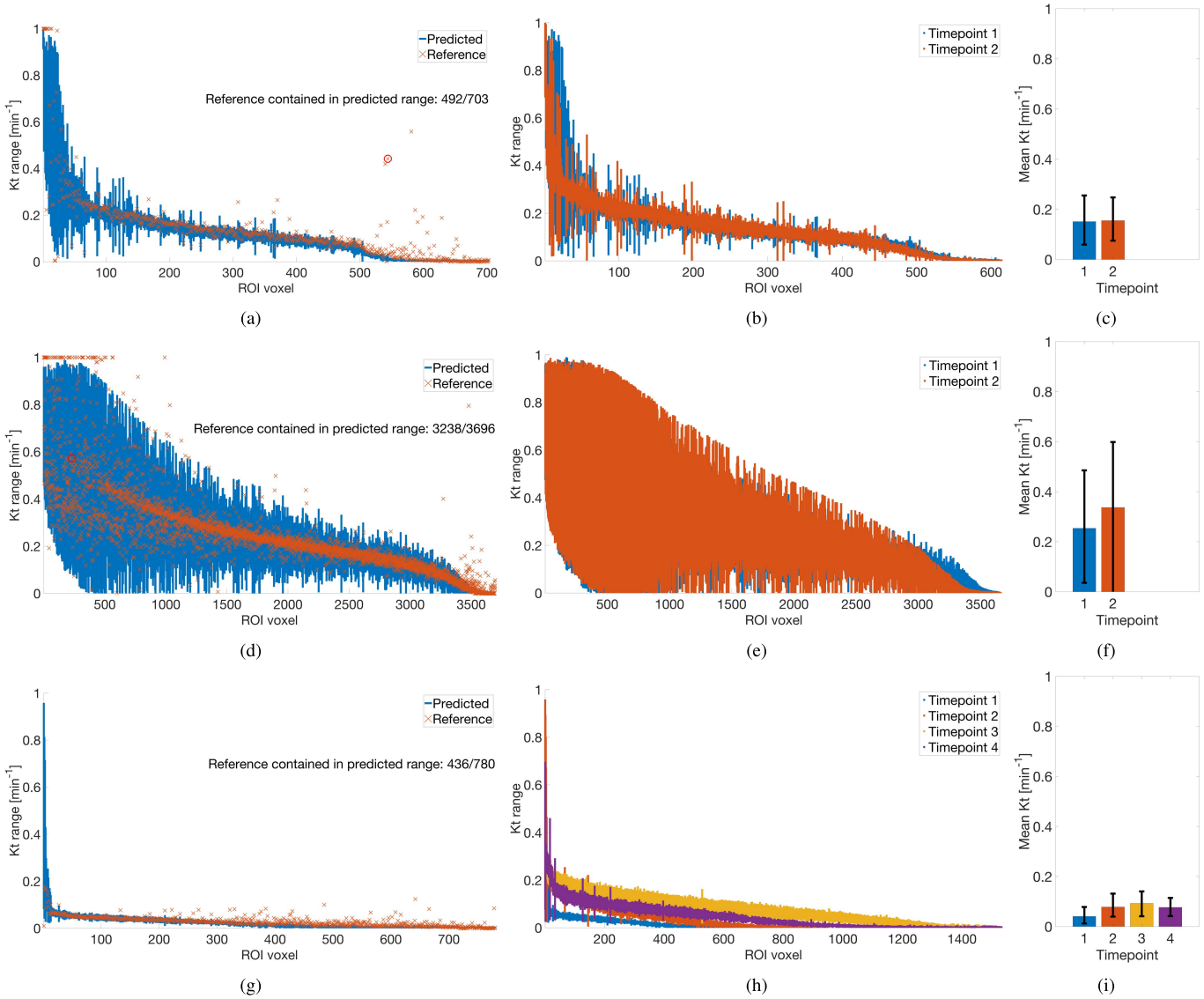

Fig. 7. Evaluation of change in K^t in tumor ROI and mean K^t over multiple time points for three lesions. Top row (a)–(c): Meningioma with no treatment and no change over time. Middle row (d)–(f): Pituitary adenoma with radiation treatment and no change over time. Bottom row (g)–(i): Metastasized melanoma treated with gamma knife which increased in size until the third exam and regressed afterwards. Left column compares predicted K^t ranges by the proposed method with results of Gauss-Newton regression. Overall good agreement of reference methods with predicted parameter ranges. Middle column shows the evolution of K^t ranges over time to be in excellent agreement with the reference diagnoses. Right column shows mean K^t of reference method in tumor ROI (bar) and upper and lower bounds on the mean K^t as predicted by the proposed approach.

methods' deployment and average or maximum bounds are applied to all subsequent measurements. Since higher-order tracer-kinetic models in DCE-MRI are non-linear, the uncertainty depends on the true parameter vector itself and global bounds may over or underestimate the actual uncertainty at the voxel level. Current DCE-MRI suffers from poor repeatability and reproducibility; and we envision that local uncertainty estimates help to sieve out and condense the reliable portion of voxel measurements.

Clinical DCE-MRI suffers from poor reproducibility overall, and a substantial source is variability from the post-processing [4], [41], [42], [56]. In Fig. 3 we attempted to mimic the multi-center study by Huang *et al.* [41], and determine if the observed poor reproducibility can in part be explained as intrinsic to the non-linear tracer-kinetic modeling. Hansen et al. [14] and Huang *et al.* [41] hypothesize voxels with low contrast uptake, either because of low permeability or because of small extra-vascular-extra-cellular space volume to leak into, to be one of the reasons for poor repeatability and reproducibility. In light of cost function cross sections, e.g., in Fig. 5a, we provide further evidence to this hypothesis. Approximately flat regions in the cost function landscape leave maximum-likelihood methods no chance but to surrender to noise and dying gradients, and to report the value they reached based on their initialization. Bayesian methods can overcome this problem by leveraging prior information. Flatness of the likelihood function causes priors to become more influential which may lead to more consistent results [14] with decreased variance in such regions. This might explain why Bayesian methods were found to outperform non-Bayesian methods

in tumor classification and grading tasks [10], in producing more accurate tracer-kinetic parameter maps [10], [14], and in achieving higher radiologist's scores [8].

Regarding the vastly different ranges of tracer-kinetic parameters for different diseases [26] it is however likely that post-processing tools including such priors need to be tailored to a specific disease and tumor type [14], and possibly end up being site specific too; which would add to poor reproducibility across sites. At least for tumors, it was found that the same priors could be reused for breast [15], prostate, and head and neck tumors [10]; whereas another brain tumor research team found priors different in parameter settings (for permeability) and functional form (for volume fractions) [8], [14].

Priors are typically designed from expert knowledge and intuition [14]. For DCE-MRI such knowledge base can only come from experience with existing post-processing tools due to lack of gold standard reference methods. Hence, these priors likely inherent problems arising due to poor repeatability and reproducibility at which point the problem becomes circular. For this reason we propose to use uninformative, universal, uniform priors. In combination with a variational approach for posterior approximation and a uniform approximate posterior distribution we arrive a problem formulation that (ideally) allows to capture the full support of tracer-kinetic parameters that could potentially explain the data while relying only on data consistency.

We show that the proposed method is able to capture and predict the entailing range of parameter estimates. A comparison to standard deviations induced by noise, however, reveals that the proposed methods tends to overestimate the observed standard deviation in the medium to high noise regime. This is likely due to a mismatch between the true posterior distribution and the assumed uniform posterior distribution. The latter is able to predict the correct support, yet does not capture the relative frequencies of the estimated values accurately. This would explain why the effect is less pronounced if uncertainty is measured in terms of range of posterior support. A remedy to this could be the use of alternative approximations such as Normal or Log-Normal assumptions. We anticipate that the same network architecture could be used for these alternative distributions in which case the predicted ellipsoid would be the covariance ellipsoid. Infinite support and fixed zero-value of the probability density at zero, respectively, would require proper treatment of potential (non-physiological) negative tracer-kinetic parameters and separate treatment of healthy tissue.

While Bayesian methods may certainly provide increased performance seen in Table. I, Fig. 2, and in [8], [10], [14], [15], many practical implementations come at the cost of increased computation time which might prohibit their use in clinical practice. Schmid *et al.* [15] report 8h on a conventional CPU and similarly sized dataset, Zhu *et al.* [17] report 23.5h on a dataset less than half the size, and Dikaios *et al.* [16] report 17h for sampling based Markov-Chain-Monte-Carlo methods which are known to be computationally demanding [57]. Tietze *et al.* [8] mention long processing time for curve fitting as one of the obstacles preventing clinical deployment, and causing curve fitting only on concentration-time-curves averaged across the tumor ROIs to still be in wide-spread use [10], [17]. For this reason we propose to use a neural network which allows Bayesian regression for a whole clinical DCE exam in less than 30s on four CPUs.

Computing voxel-wise variance estimates in addition to parameter estimates raises the question of how to incorporate the added information into clinical decision making and patient management. There are many possibilities that require further investigation in collaboration with patient management teams. The spectrum of possibilities includes simply displaying parameter and uncertainty maps side-by-side, to discarding high variance voxels, or reporting a single representative biomarker and its error bound. In Fig. 7 we illustrate two possible ways of data distillation and representation. While the temporal evolution of predicted K^t ranges matches the reference diagnosis accurately, further compression of the data into a single biomarker was not successful in providing desired specificity.

Although care must be taken because of differences in the tumor type under investigation, when comparing similar mean K^t, this relatively wide spread of mean K^t parameter in the tumor ROI is also observed by Huang et al. [41] where the mean K^t in the tumor ROI was computed with different software tools at different sites. Because baseline T_{10} and arterial input function were fixed in [41], the variation in mean K^t observed across different regression tools could be intrinsic to the assumed tracer-kinetic modeling. Hence, this initial study could hint toward reasons for current challenges in the development of reliable biomarkers [35], [51]. Yet, further investigation on larger patient cohorts is crucial to better determine how uncertainty estimates can help to inform clinical decision making.

The proposed network performed well on the simulation test data and clinical data and in comparison with non-Bayesian regression, which is in line with other works on this topic [8], [10], [14], [15]. However, there were some voxel cases in the clinical dataset for which the predicted parameter differed substantially from the reference method (Fig. 7). A closer look in Fig. 5a reveals this to be due to high noise and non-convexity of the cost function which could make multiple settings of parameters equally plausible. The effects of this problem have been noticed in previous work [14], [41]. Although clinical consequences remain unclear [14], [41], the proposed method in (9) presents a way of addressing the problem by consistently enforcing the physiologically realistic solution through memorization in the neural network weights.

In this work we used artificial training concentration-time-curves generated with arterial input functions extracted from real clinical DCE-MRI exam. In principle the proposed method is not limited to be trained with artificial data but could also be trained with concentration-time-curves from clinical exams. One obstacle is data balance which is known to be important for training neural networks. For artificial data such balance can readily be achieved by choice of data generation. For clinical data, the frequency of occurrence of tracer-kinetic parameters follows their natural distribution. This gives a highly skewed dataset in which high values of K^t for example appear less often. Hence, the network observes such values less often during training, which might

cause degradation in performance for such parameter values. Although less frequent, such high enhancement regions are known to be important for diagnosis and accuracy should not be compromised. Successful design of training data consisting of real clinical concentration-time-curves is subject to future research.

The proposed method could be applied to other quantitative imaging techniques such as Dynamic Position Emission Tomography (PET) [17], [58], [59], Dynamic-Susceptibility-Contrast (DSC) MRI [60], [61], and Arterial Spin Labeled (ASL) MRI [19]. The proposed method is also not limited to Brain nor the deployed extended Tofts-Kety model. Other possible areas of application include Prostate, Breast, Liver, and Heart; and include onset time and delay estimation [6], [10], [14] as well as higher-order tracer kinetic models [1], which share the non-strictly-convex characteristics of the cost function landscape [6], [14], [25]. Model selection itself is one of the prominent problems in DCE-MRI [62]. An extension of the proposed method would be to incorporate model selection into the estimation of uncertainty via the introduction of hyperpriors and automatic relevance determination [63].

Acknowledgment

The authors would like to thank D. Hwang for help identifying, anonymizing, and processing clinical data. They also like to thank members of the USC Brain Tumor Board for feedback and discussion.

References

[1] S. P. Sourbron and D. L. Buckley, "Classic models for dynamic contrast-enhanced MRI," *NMR Biomed.*, vol. 26, no. 8, pp. 1004–1027, May 2013.

[2] B. F. Kurland *et al.*, "Promise and pitfalls of quantitative imaging in oncology clinical trials," *Magn. Reson. Imag.*, vol. 30, no. 9, pp. 1301–1312, Nov. 2012.

[3] Y. Cao, "The promise of dynamic contrast-enhanced imaging in radiation therapy," *Seminars Radiat. Oncol.*, vol. 21, no. 2, pp. 147–156, Apr. 2011.

[4] W. Huang *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*, vol. 2, no. 1, pp. 56–66, Mar. 2016.

[5] B. M. Kelm, B. H. Menze, O. Nix, C. Zechmann, and F. A. Hamprecht, "Estimating kinetic parameter maps from dynamic contrast-enhanced MRI using spatial prior knowledge," *IEEE Trans. Med. Imag.*, vol. 28, no. 10, pp. 1534–1547, Oct. 2009.

[6] M. R. Orton *et al.*, "Bayesian estimation of pharmacokinetic parameters for DCE-MRI with a robust treatment of enhancement onset time," *Phys. Med. Biol.*, vol. 52, no. 9, pp. 2393–2408, Apr. 2007.

[7] R. G. P. Lopata, W. H. Backes, P. P. J. van den Bosch, and N. A. W. van Riel, "On the identifiability of pharmacokinetic parameters in dynamic contrast-enhanced imaging," *Magn. Reson. Med.*, vol. 58, no. 2, pp. 425–429, Jul. 2007.

[8] A. Tietze *et al.*, "Bayesian modeling of dynamic contrast enhanced MRI data in cerebral glioma patients improves the diagnostic quality of hemodynamic parameter maps," *PLoS ONE*, vol. 13, no. 9, pp. 1–14, Sep. 2018.

[9] T. S. Ahearn, R. T. Staff, T. W. Redpath, and S. I. K. Semple, "The use of the Levenberg–Marquardt curve-fitting algorithm in pharmacokinetic modelling of DCE-MRI data," *Phys. Med. Biol.*, vol. 50, no. 9, p. N85, Apr. 2005.

[10] N. Dikaios *et al.*, "A comparison of Bayesian and non-linear regression methods for robust estimation of pharmacokinetics in DCE-MRI and how it affects cancer diagnosis," *Comput. Med. Imag. Graph.*, vol. 56, pp. 1–10, Mar. 2017.

[11] C. Ulas *et al.*, "Convolutional neural networks for direct inference of pharmacokinetic parameters: Application to stroke dynamic contrast-enhanced MRI," *Frontiers Neurology*, vol. 9, p. 1147, Jan. 2019.

[12] A. Garpebring *et al.*, "Uncertainty estimation in dynamic contrast-enhanced MRI," *Magn. Reson. Med.*, vol. 69, no. 4, pp. 992–1002, Apr. 2013.

[13] K. Mouridsen, M. B. Hansen, L. Østergaard, and S. N. Jespersen, "Reliable estimation of capillary transit time distributions using DSC-MRI," *J. Cerebral Blood Flow Metabolism*, vol. 34, no. 9, pp. 1511–1521, 2014.

[14] M. B. Hansen *et al.*, "Robust estimation of hemo-dynamic parameters in traditional DCE-MRI models," *PLoS ONE*, vol. 14, no. 1, 2019, Art. no. e0209891.

[15] V. J. Schmid, B. Whitcher, A. R. Padhani, N. J. Taylor, and G.-Z. Yang, "Bayesian methods for pharmacokinetic models in dynamic contrast-enhanced magnetic resonance imaging," *IEEE Trans. Med. Imag.*, vol. 25, no. 12, pp. 1627–1636, Dec. 2006.

[16] N. Dikaios, S. Arridge, V. Hamy, S. Punwani, and D. Atkinson, "Direct parametric reconstruction from undersampled (k, t)-space data in dynamic contrast enhanced MRI," *Med. Image Anal.*, vol. 18, no. 7, pp. 989–1001, 2014.

[17] W. Zhu *et al.*, "A Bayesian spatial temporal mixtures approach to kinetic parametric images in dynamic positron emission tomography," *Med. Phys.*, vol. 43, no. 3, pp. 1222–1234, Mar. 2016.

[18] A. Sitek, Q. Li, G. El Fakhri, and N. M. Alpert, "Validation of Bayesian analysis of compartmental kinetic models in medical imaging," *Phys. Medica*, vol. 32, no. 10, pp. 1252–1258, Oct. 2016.

[19] M. A. Chappell, A. R. Groves, B. Whitcher, and M. W. Woolrich, "Variational Bayesian inference for a nonlinear forward model," *IEEE Trans. Signal Process.*, vol. 57, no. 1, pp. 223–236, Jan. 2009.

[20] E. T. Jaynes, "Information theory and statistical mechanics," *Phys. Rev.*, vol. 106, no. 4, pp. 620–630, 1957.

[21] Y. Guo, S. G. Lingala, Y. Zhu, R. M. Lebel, and K. S. Nayak, "Direct estimation of tracer-kinetic parameter maps from highly undersampled brain dynamic contrast enhanced MRI," *Magn. Reson. Med.*, vol. 78, no. 4, pp. 1566–1578, 2017.

[22] F. F. J. Simonis, A. Sbrizzi, E. Beld, J. J. W. Lagendijk, and C. A. T. van den Berg, "Improving the arterial input function in dynamic contrast enhanced MRI by fitting the signal in the complex plane," *Magn. Reson. Med.*, vol. 76, no. 4, pp. 1236–1245, Oct. 2016.

[23] D. De Naeyer, Y. De Deene, W. P. Ceelen, P. Segers, and P. Verdonck, "Precision analysis of kinetic modelling estimates in dynamic contrast enhanced MRI," *Magn. Reson. Mater. Phys., Biol. Med.*, vol. 24, no. 2, pp. 51–66, 2011.

[24] DCE MRI Technical Committee, "DCE MRI quantification profile," Radiological Society of North America Quantitative Imaging Biomarkers Alliance. Version 1.0. Reviewed Draft, Jul. 2012.

[25] S. P. Sourbron and D. L. Buckley, "On the scope and interpretation of the Tofts models for DCE-MRI," *Magn. Reson. Med.*, vol. 66, no. 3, pp. 735–745, Mar. 2011.

[26] A. K. Heye, R. D. Culling, M. del C. Valdés Hernández, M. J. Thrippleton, and J. M. Wardlaw, "Assessment of blood–brain barrier disruption using dynamic contrast-enhanced MRI. A systematic review," *NeuroImage, Clin.*, vol. 6, pp. 262–274, Jan. 2014.

[27] M. C. Schabel, "A unified impulse response model for DCE-MRI," *Magn. Reson. Med.*, vol. 68, no. 5, pp. 1632–1646, Jan. 2012.

[28] A. V. Dalca, G. Balakrishnan, J. Guttag, and M. R. Sabuncu, "Unsupervised learning for fast probabilistic diffeomorphic registration," in *Proc. Med. Image Comput. Comput. Assist. Intervent. (MICCAI)*, in Lecture Notes in Computer Science, vol. 11070, 2018, pp. 729–738.

[29] M. I. Jordan and M. J. Wainwright, "Graphical models, exponential families, and variational inference," *Found. Trends Mach. Learn.*, vol. 1, nos. 1–2, pp. 1–305, 2007.

[30] G. A. Darbellay and I. Vajda, "Entropy expressions for multivariate continuous distributions," *IEEE Trans. Inf. Theory*, vol. 46, no. 2, pp. 709–712, Mar. 2000.

[31] G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *Science*, vol. 313, no. 5786, pp. 504–507, 2006.

[32] D. P. Kingma and M. Welling, "Auto-encoding variational Bayes," Dec. 2013, *arXiv:1312.6114*. [Online]. Available: https://arxiv.org/abs/1312.6114

[33] C. A. Cuenod and D. Balvay, "Perfusion and vascular permeability: Basic concepts and measurement in DCE-CT and DCE-MRI," *Diagnostic Interventional Imag.*, vol. 94, no. 12, pp. 1187–1204, Dec. 2013.

[34] C. Lavini, "Simulating the effect of input errors on the accuracy of Tofts' pharmacokinetic model parameters," *Magn. Reson. Imag.*, vol. 33, no. 2, pp. 222–235, Feb. 2015.

[35] R. A. Little *et al.*, "Evaluation of dynamic contrast-enhanced MRI biomarkers for stratified cancer medicine: How do permeability and perfusion vary between human tumours," *Magn. Reson. Imag.*, vol. 46, pp. 98–105, Feb. 2018.

[36] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*. [Online]. Available: https://arxiv.org/abs/1412.6980

[37] R. J. Bosca and E. F. Jackson, "Creating an anthropomorphic digital MR phantom—An extensible tool for comparing and evaluating quantitative imaging algorithms," *Phys. Med. Biol.*, vol. 61, no. 2, pp. 974–982, Jan. 2016.

[38] E. Henderson, J. Sykes, D. Drost, H.-J. Weinmann, B. K. Rutt, and T.-Y. Lee, "Simultaneous MRI measurement of blood flow, blood volume, and capillary permeability in mammary tumors using two different contrast agents," *J. Magn. Reson. Imag.*, vol. 12, no. 6, pp. 991–1003, Nov. 2000.

[39] D. L. Buckley, "Uncertainty in the analysis of tracer kinetics using dynamic contrast-enhanced T_1-weighted MRI," *Magn. Reson. Med.*, vol. 47, no. 3, pp. 601–606, Feb. 2002.

[40] K. Murase, "Efficient method for calculating kinetic parameters using T_1-weighted dynamic contrast-enhanced magnetic resonance imaging," *Magn. Reson. Med.*, vol. 51, no. 4, pp. 858–862, Mar. 2004.

[41] W. Huang *et al.*, "Variations of dynamic contrast-enhanced magnetic resonance imaging in evaluation of breast cancer therapy response: A multicenter data analysis challenge," *Transl. Oncol.*, vol. 7, no. 1, pp. 153–166, Feb. 2014.

[42] L. Beuzit *et al.*, "Dynamic contrast-enhanced MRI: Study of intersoftware accuracy and reproducibility using simulated and clinical data," *J. Magn. Reson. Imag.*, vol. 43, no. 6, pp. 1288–1300, Dec. 2015.

[43] S. C. L. Deoni, T. M. Peters, and B. K. Rutt, "High-resolution T_1 and T_2 mapping of the brain in a clinically acceptable time with DESPOT1 and DESPOT2," *Magn. Reson. Med.*, vol. 53, no. 1, pp. 237–241, Dec. 2005.

[44] B. M. Ellingson, P. Y. Wen, and T. F. Cloughesy, "Modified criteria for radiographic response assessment in glioblastoma clinical trials," *Neurotherapeutics*, vol. 14, no. 2, pp. 307–320, Apr. 2017.

[45] D. Shanbhag, S. N. Gupta, K. T. Rajamani, Y. Zhu, and R. Mullick, "A generalized methodology for detection of vascular input function with dynamic contrast enhanced perfusion data," in *Proc. Intl. Soc. Mag. Reson. Med.*, vol. 12, 2012, p. 10.

[46] S. L. S. Chan and Y. Gal, "Automatic detection of arterial Voxels in dynamic contrast-enhanced mr images of the brain," in *Proc. Int. Conf. Digit. Image Comput. Techn. Appl. (DICTA)*, Dec. 2012, pp. 1–7.

[47] M. Jacobs, M. Benovoy, L.-C. Chang, A. E. Arai, and L.-Y. Hsu, "Evaluation of an automated method for arterial input function detection for first-pass myocardial perfusion cardiovascular magnetic resonance," *J. Cardiovascular Magn. Reson.*, vol. 18, no. 1, p. 17, Apr. 2016.

[48] H. Lu, C. Clingman, X. Golay, and P. C. M. Van Zijl, "Determining the longitudinal relaxation time (T_1) of blood at 3.0 tesla," *Magn. Reson. Med.*, vol. 52, no. 3, pp. 679–682, Aug. 2004.

[49] G. J. Parker *et al.*, "Experimentally-derived functional form for a population-averaged high-temporal-resolution arterial input function for dynamic contrast-enhanced MRI," *Magn. Reson. Med.*, vol. 56, no. 5, pp. 993–1000, 2006.

[50] R. M. Lebel *et al.*, "The comprehensive contrast-enhanced neuro exam," in *Proc. 23th Sci. Meeting, Int. Soc. Magn. Reson. Med.*, Toronto, CN, Canada, 2015, p. 3705.

[51] N. Just, "Improving tumour heterogeneity MRI assessment with histograms," *Brit. J. Cancer*, vol. 111, no. 12, pp. 2205–2213, Sep. 2014.

[52] R. Venkatasubramanian, R. B. Arenas, M. A. Henson, and N. S. Forbes, "Mechanistic modelling of dynamic MRI data predicts that tumour heterogeneity decreases therapeutic response," *Brit. J. Cancer*, vol. 103, no. 4, pp. 486–497, Jul. 2010.

[53] Y.-C. Chang, C.-S. Huang, Y.-J. Liu, J.-H. Chen, Y.-S. Lu, and W.-Y. I. Tseng, "Angiogenic response of locally advanced breast cancer to neoadjuvant chemotherapy evaluated with parametric histogram from dynamic contrast-enhanced MRI," *Phys. Med. Biol.*, vol. 49, no. 16, pp. 3593–3602, Aug. 2004.

[54] S. L. Peng *et al.*, "Analysis of parametric histogram from dynamic contrast-enhanced MRI: Application in evaluating brain tumor response to radiotherapy," *NMR Biomed.*, vol. 26, no. 4, pp. 443–450, Oct. 2012.

[55] C. Hayes, A. R. Padhani, and M. O. Leach, "Assessing changes in tumour vascular function using dynamic contrast-enhanced magnetic resonance imaging," *NMR Biomed.*, vol. 15, no. 2, pp. 154–163, Feb. 2002.

[56] T. Heye *et al.*, "Reproducibility of dynamic contrast-enhanced mr imaging—Part I. Perfusion characteristics in the female pelvis by using multiple computer-aided diagnosis perfusion analysis solutions," *Radiology*, vol. 266, no. 3, pp. 801–811, Mar. 2013.

[57] C. M. Bishop, *Pattern Recognition and Machine Learning (Information Science and Statistics)*. New York, NY, USA: Springer-Verlag, 2006.

[58] F. A. Kotasidis, C. Tsoumpas, and A. Rahmim, "Advanced kinetic modelling strategies: Towards adoption in clinical PET imaging," *Clin. Transl. Imag.*, vol. 2, no. 3, pp. 219–237, Jun. 2014.

[59] J. Yan, B. Planeta-Wilson, and R. E. Carson, "Direct 4-D PET list mode parametric reconstruction with a novel EM algorithm," *IEEE Trans. Med. Imag.*, vol. 31, no. 12, pp. 2213–2223, Dec. 2012.

[60] K. Mouridsen, K. Friston, N. Hjort, L. Gyldensted, L. Østergaard, and S. Kiebel, "Bayesian estimation of cerebral perfusion using a physiological model of microvasculature," *NeuroImage*, vol. 33, no. 2, pp. 570–579, Nov. 2006.

[61] A. Mehndiratta, F. Calamante, B. J. Macintosh, D. E. Crane, S. J. Payne, and M. A. Chappell, "Modeling and correction of bolus dispersion effects in dynamic susceptibility contrast MRI," *Magn. Reson. Med.*, vol. 72, no. 6, pp. 1762–1774, Jan. 2014.

[62] C. Duan, J. F. Kallehauge, G. L. Bretthorst, K. Tanderup, J. J. H. Ackerman, and J. R. Garbow, "Are complex DCE-MRI models supported by clinical data," *Magn. Reson. Med.*, vol. 77, no. 3, pp. 1329–1339, Mar. 2016.

[63] T. E. J. Behrens, H. J. Berg, S. Jbabdi, M. F. S. Rushworth, and M. W. Woolrich, "Probabilistic diffusion tractography with multiple fibre orientations: What can we gain," *NeuroImage*, vol. 34, no. 1, pp. 144–155, Jan. 2007.