FULL PAPER

A model selection framework to quantify microvascular liver function in gadoxetate-enhanced MRI: Application to healthy liver, diseased tissue, and hepatocellular carcinoma

Michael Berks[1] | Ross A. Little[1] | Yvonne Watson[1] | Sue Cheung[1] |
Anubhav Datta[1,2] | James P. B. O'Connor[1,2,3] | Davide Scaramuzza[4] | Geoff J. M. Parker[1,5,6]

[1]Division of Cancer Sciences, Quantitative Biomedical Imaging Laboratory, University of Manchester, Manchester, UK

[2]The Christie NHS Foundation Trust, Manchester, UK

[3]Division of Radiotherapy and Imaging, Institute of Cancer Research, London, UK

[4]Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy

[5]Bioxydyn Ltd, Manchester, UK

[6]Centre for Medical Image Computing, University College London, London, UK

Correspondence
Michael Berks, Division of Cancer Sciences, Quantitative Biomedical Imaging Laboratory, University of Manchester, Manchester, M13 9PT, UK.
Email: michael.berks@manchester.ac.uk

Funding information
Bayer, Grant/Award Number: 104/14; CRUK and EPSRC Cancer Imaging Centre in Cambridge and Manchester, Grant/Award Number: C8742/A18097; Cancer Research UK, Grant/Award Number: C19221/A22746

Purpose: We introduce a novel, generalized tracer kinetic model selection framework to quantify microvascular characteristics of liver and tumor tissue in gadoxetate-enhanced dynamic contrast-enhanced MRI (DCE-MRI).

Methods: Our framework includes a hierarchy of nested models, from which physiological parameters are derived in 2 regimes, corresponding to the active transport and free diffusion of gadoxetate. We use simulations to show the sensitivity of model selection and parameter estimation to temporal resolution, time-series duration, and noise. We apply the framework in 8 healthy volunteers (time-series duration up to 24 minutes) and 10 patients with hepatocellular carcinoma (6 minutes).

Results: The active transport regime is preferred in 98.6% of voxels in volunteers, 82.1% of patients' non-tumorous liver, and 32.2% of tumor voxels. Interpatient variations correspond to known co-morbidities. Simulations suggest both datasets have sufficient temporal resolution and signal-to-noise ratio, while patient data would be improved by using a time-series duration of at least 12 minutes.

Conclusions: In patient data, gadoxetate exhibits different kinetics: (a) between liver and tumor regions and (b) within regions due to liver disease and/or tumor heterogeneity. Our generalized framework selects a physiological interpretation at each voxel, without preselecting a model for each region or duplicating time-consuming optimizations for models with identical functional forms.

KEYWORDS
gadoxetate, hepatocellular carcinoma, model selection, quantitative DCE-MRI, tracer kinetic modeling

1 | INTRODUCTION

1.1 | Gadoxetate-enhanced MR imaging of the liver

Gadoxetate disodium (Eovist or Primovist, Bayer Healthcare, Berlin, Germany) is a gadolinium-based hepatobiliary contrast agent taken up by hepatocytes and excreted by the biliary pathway, allowing more direct measurement of liver function than standard extracellular contrast agents. It is regularly used in the clinical assessment of chronic liver disease and cancer.[1-4] In recent meta-analyses of hepatocellular carcinoma (HCC, most common primary malignancy of the liver[5] and second leading cause of cancer mortality worldwide[6]) detection, gadoxetate-enhanced MRI returned the highest overall per-lesion sensitivity and positive predictive value, compared with contrast-enhanced ultra-sound, CT, and MRI with extracellular contrast agents.[7,8]

In addition to diagnosing HCC, relative lesion intensity to surrounding liver tissue has been shown to characterize subtypes of the tumor, indicating progression of nodules to hypervascular HCC,[9] microvascular invasion,[10] and correlating with prognostic histological biomarkers.[11] Gadoxetate-enhanced imaging may also help identify and grade chronic liver disease, including portal hypertension[12] and fibrosis,[13] often co-present in patients with HCC.

1.2 | Tracer kinetic modeling in DCE-MRI of the liver

Quantitative dynamic contrast enhanced (DCE) MRI, where tracer kinetic models are fitted to image time series, enables estimating parameters reflecting tissue microvascular function. In metastatic liver tumors, there is a long history of tracer kinetic modeling using standard gadolinium-based contrast agents.[14-17] A common approach assumes metastases recruit an arterial-only supply, with passive exchange of contrast agent between separately measurable plasma and interstitial spaces, enabling the fit of a 2-compartment extended-Tofts model.[18] Model parameters, and in particular, the intercompartment volume transfer constant K^{trans}, can provide clinically useful endpoints in trials of antivascular drugs.[14,17]

To account for liver parenchyma's arterial *and* hepatic portal vein blood supply, a dual-input single-compartment model was proposed for application outside tumors,[19] and used in patient studies showing a correlation between perfusion parameters and the severity of cirrhosis or fibrosis.[20] The single-compartment model was generalized into a 2-compartment model suitable for parameterizing tracer kinetics in liver tissue and neuroendocrine metastatic lesions.[21,22] In an alternative approach, the dual-input single-compartment model and an extended-Tofts model were fitted to metastatic cancers,[15]

and the Aikake information criterion (AIC)[23] used to select the most appropriate model at each voxel.

There have been far fewer quantitative DCE-MRI studies of HCC than metastatic lesions: applying an extended-Tofts model, decreased K^{trans} on first scan has been shown to predict survival,[24] and the dual-input single-compartment model used to quantify HCC perfusion.[25] In the only study we are aware of applying quantitative DCE-MRI to gadoxetate enhanced HCC tumors,[26] fitting a single-input (arterial only) Tofts model showed correlations for K^{trans} and the extracellular, extravascular volume v_e to histological grade and tumor microvascular density, respectively.

Other tracer kinetic modeling studies of gadoxetate include initial attempts to estimate intracellular uptake fraction as a marker of liver function[27,28]; however, these were limited by using models comprising a single-input blood supply, which further failed to account for active transport of gadoxetate into the intracellular space. These shortcomings were resolved in a dual-input, uptake model[29,30] that provided significantly better fit than the single-compartment model. To model gadoxetate dynamics through to the hepatobiliary phase (typically reached 20 minutes post-injection), an efflux term was added, and applied to time series acquired using an 50-minute protocol in healthy volunteers.[31] Recently, a 3-compartment model was proposed to quantify hepatic perfusion and hepatocyte function in patients with chronic liver disease, applied to a 38-minute acquisition.[13]

1.3 | Modeling challenges

With limited previous studies, and different models applied in each study, there is no common method for applying tracer kinetic modeling to gadoxetate-enhanced images. There are several challenges to consider: contrast dynamics in tissue where hepatocytes actively transport gadoxetate will be very different to tumor cells where gadoxetate behaves more like a passively transported agent, indicating the need for (at least) 2 forms of tracer-kinetic model. Moreover, in tumors native to the liver such as HCC, there may be residual uptake of contrast agent through the organic anion polypeptide (OATP) pathway due to varying OATP function in the cells of different tumor subtypes,[1] meaning that even with well-defined tumor regions-of-interest (ROIs), it is not certain *a priori* which model will be appropriate in any voxel. Thus computing parameter estimates over a whole ROI risks including voxels that do not meet a model's physiological assumptions, potentially leading to unreliable estimates and misleading interpretations. This suggests voxel-wise model selection may be a more suitable method for analyzing the data.[15,32] Common co-morbidities associated with HCC provide additional complications, with patients having a range of liver

function and signs of cirrhosis, fibrosis and splenomegaly which may further confound model assumptions if analysis is extended to the non-tumorous liver.

We investigate fitting tracer kinetic models to individual voxels of gadoxetate-enhanced MRI sequences, using models developed to explain the active transport of gadoxetate and the 2-compartment exchange model (and its derivatives including the Tofts model).[18,19,29,31,33] We introduce a novel model selection framework in which previously applied models can be derived as specific instances of generalized functional forms. In doing so, we highlight mathematical connections between the models, using these to hypothesize which will better fit different tissue types. We test these hypothesizes using extensive Monte Carlo simulations, exploring model selection criteria. We compute expectations on the precision and accuracy of derived model parameters, and show how these are affected by sequence duration, temporal resolution, and signal noise. Finally, we test our hypothesizes on 2 in vivo datasets: the healthy volunteers presented in Ref. [31] and a new dataset of 10 patients with HCC, showing links between model selection and patient co-morbidities, determining the limitations of the data and using these to make recommendations for designing further patient studies.

2 | METHODS

## 2.1 | Tracer-kinetic modeling

2.1.1 | Tissue models

Georgiou et al[31] presented a model to describe both the active uptake of gadoxetate contrast agent by hepatocytes and the subsequent efflux of the agent into the bile duct (Figure 1A). The model incorporates a dual-input function to allow for arterial and hepatic portal vein blood supply, convolved with a biexponential impulse response function (IRF):

$$C^{AUEM}(t) = F_p \left[E_+ e^{-tT_i^{-1}} + \left(1 - E_+\right) e^{-tT_e^{-1}} \right] * C_p(t) \quad (1)$$

$$E_+ = \frac{E_i}{1 - T_e/T_i} \quad (2)$$

$$E_i = \frac{K_i}{F_p + K_i} \quad T_e = \frac{v_{ecs}}{F_p + K_i} \quad T_i = \frac{v_i}{K_{ef}} \quad (3)$$

$$C_p(t) = f_a C_a(t - \tau_a) + (1 - f_a) C_v(t - \tau_v) \quad (4)$$
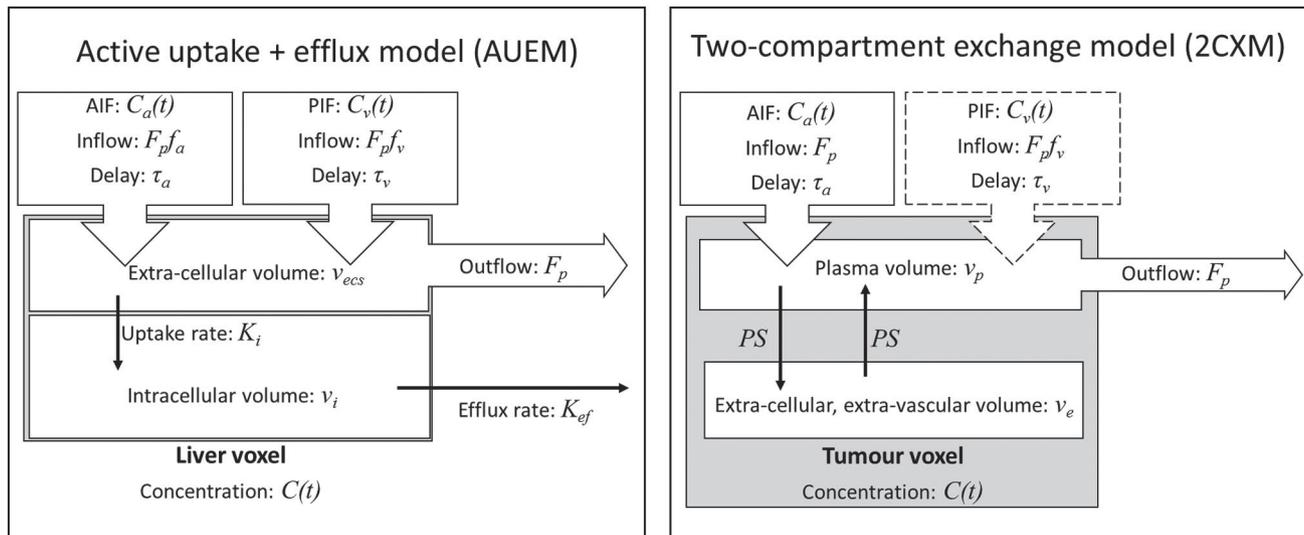
FIGURE 1 DCE model diagrams A, The active uptake + efflux model (AUEM) is a 2-compartment biexponential model with dual arterial and portal vein plasma input, designed to model the active transport of gadoxetate contrast agent by hepatocytes. Contrast agent is assumed to be instantaneously well mixed in the extracellular volume v_{ecs}, and then actively transported into the intracellular volume v_i with uptake rate K_i, which together comprise the whole voxel space. Gadoxetate is later excreted to the bile ducts at rate K_{ef}. The model can be simplified first assuming zero efflux ($K_{ef} = 0$), then zero uptake $K_i = 0$, which reduces to the dual-input single compartment model. B, In the 2-compartment exchange model (2CXM), the interstitial v_e and plasma volumes v_p are assumed separable, with measurable passive-exchange between them at a rate depending on the permeability surface area product PS. It is assumed cells in these voxels do not actively uptake contrast agent, and thus there is no contribution to signal change from the intracellular space. See Sourbron and Buckley[34] for a comprehensive discussion of simplifications of the 2CXM to the other passive-exchange regime models such as the (extended-) Tofts model. The 2CXM is often considered with a single arterial input; however, our framework includes a portal vein supply (dashed box), generating a dual-input model with the same functional form as the AUEM

where C_a and C_v are the arterial and venous input functions of plasma concentration, offset by τ_a(minutes) and τ_v(minutes) and $f_a \in [0, 1]$ controls the proportion of each input. The IRF is parameterized by: uptake rate K_i(min^{-1}); efflux rate, K_{ef}(min^{-1}); plasma flow, F_p(mlmin^{-1}ml^{-1}), and extracellular volume v_{ecs}(mlml^{-1}). The model assumes v_{ecs} comprises the plasma volume (v_p) and interstitial space v_e, with rapid exchange between the 2 (ie,, endothelial permeability $\to \infty$) that immediately reaches equilibrium, such that v_{ecs} acts as a single compartment.[29,34] In addition, it is assumed the volume of non-hepatocyte cells in the intracellular space v_i is negligible, so that $v_i = 1 - v_{ecs}$, and the contribution of contrast agent in bile canaliculi can be ignored.[35,36]

We label this the active uptake + efflux model (AUEM), noting it extended an earlier uptake-only model described by Sourbron et al,[29] derived from the full AUEM by setting the efflux rate $K_{ef} = 0$. Setting the uptake rate $K_i = 0$, further reduces the uptake-only model to a dual-input single-compartment model.[19]

This hierarchy of nested models evokes the analysis by Sourbron and Buckley,[34] in which it is shown a biexponential 2-compartment exchange model (2CXM, Figure 1B), with physiological parameters F_p (plasma flow), PS (permeability surface area product, min^{-1}), v_e, and v_p, reduces to the extended-Tofts (assuming $F_p = \infty$) and Tofts model (assuming $v_p = 0$).

In a similar vein, we note that while the 2CXM is usually considered with a single input function, the form of its IRF is identical to the AUEM, and thus if we fit a dual-input 2CXM, with no restrictions on the parameters, the modeled contrast time series is be identical to the AUEM—it is only the derivation and interpretation of the physiological parameters that differ.

We hypothesize that the AUEM (or one of its simplifications) suitably describes the dynamics of gadoxetate in healthy liver tissue, whereas a 2CXM is more likely to fit tumor voxels, where hepatocytes are not present to actively transport gadoxetate, and the plasma and interstitial spaces are separable volumes v_p and v_e, with a measurable bisymmetric exchange of gadoxetate between them.

This leads us to adopt the following approach: we fit as the most general model, a dual-input biexponential of the form

$$C^{I_4}(t) = \left[\alpha_+ e^{-t\beta_+} + \alpha_- e^{-t\beta_-}\right] * C_p(t) \quad (5)$$

This has 6 free parameters to optimize: the IRF functional parameters $\alpha_\pm$, $\beta_\pm$, and physiologically meaningful parameters controlling the arterial input fraction and delay time f_a and $\tau_a = \tau_v$ (Equation (4), see also Supporting Information: Delay parameters). We then derive 4 further physiological parameters from the functional parameters, using either the active uptake or passive exchange interpretations (see Supporting Information: Model derivations for full complete intermediary equations):

2.2 | Active uptake (F_p, $v_{ecs} = 1 - v_i$, K_i, K_{ef})

$$F_p = \alpha_+ + \alpha_- \quad (6)$$

$$v_{ecs} = \frac{(\alpha_+ + \alpha_-)^2}{\alpha_+\beta_+ + \alpha_-\beta_-} \quad (7)$$

$$K_i = \frac{\alpha_+(\alpha_+ + \alpha_-)(\beta_- - \beta_+)}{\alpha_+\beta_+ + \alpha_-\beta_-} \quad (8)$$

$$K_{ef} = \beta_+ \left[1 - \frac{(\alpha_+ + \alpha_-)^2}{\alpha_+\beta_+ + \alpha_-\beta_-}\right] \quad (9)$$

2.3 | Passive exchange (F_p, PS, v_e, v_p)

$$F_p = \alpha_+ + \alpha_- \quad (10)$$

$$PS = \frac{\alpha_+\alpha_-(\alpha_+ + \alpha_-)(\beta_+ - \beta_-)^2}{(\alpha_+\beta_+ + \alpha_-\beta_-)^2} \quad (11)$$

$$v_e = \frac{\alpha_+\alpha_-(\beta_+ - \beta_-)^2}{\beta_+\beta_-(\alpha_+\beta_+ + \alpha_-\beta_-)} \quad (12)$$

$$v_p = \frac{(\alpha_+ + \alpha_-)^2}{\alpha_+\beta_+ + \alpha_-\beta_-} \quad (13)$$

Thus, the 2 model regimes share 3 physiological parameters (F_p, f_a, and τ_a), differing only in the interpretation of the 2 compartment volumes and the transfer rates. Moreover, although they differ in physiological interpretation, the extracellular volume v_{ecs} in the active-uptake regime is *functionally* equivalent to the vascular volume v_p in the 2CXM (both being the initial compartment receiving vascular input; Equations (7) and (13)).

These interpretations impose constraints on valid ranges for the parameters—in the active-uptake regime $v_{ecs} \leq 1$ and in the passive-exchange regime, $v_e + v_b \leq 1$ (where $v_b = v_p/(1 - \text{Hct})$). However, we do not impose these constraints during model fitting. Instead, the 4 functional parameters $\alpha_\pm$, $\beta_\pm$ are simply constrained to be non-negative, and then post-fit, we analyze whether there is a valid physiological interpretation of the parameters under either regime. This allows the optimizer to explore the complete parameter space of *both* physiological models without requiring a prior assumption of which specific instance should be preferred.

2.4 | Nested models

In addition to the full biexponential model, we fit a hierarchy of progressively simplified models, setting first $\beta_+ = 0$,

then $\alpha_+ = 0$. The nested models are defined by the form of their impulse responses, taking 4, 3, and 2 parameters, respectively, labeled as the I_4, I_3, and I_2 models. When we derive physiological parameters in either the active-uptake or passive-exchange regimes from a nested model, we obtain parameters that match one of the reduced forms of the AUEM or 2CXM previously outlined in **Tissue models**. We present these specific forms below.

2.5 | I_3 model:
$$C^{I_3}(t) = \left[\alpha_+ + \alpha_- e^{-t\beta_-}\right] * C_p(t), \quad \beta_+ = 0$$

In the active-uptake regime, when $Bpos = 0$ from Equations (6)-(9), we see that F_p, v_{ecs}, and K_i are well-defined and positive, while $K_{ef} = 0$. Thus, the I_3 model is functionally equivalent to the uptake-only model.[29] In the passive-exchange regime, v_{ecs} is ill-defined, while F_p, PS, and v_p are positive. This is described as the 2-compartment uptake model in Ref. [33], and may be appropriate in tissues where backflow of the tracer into the plasma space is negligible, where there are poorly mixed compartments, or extraction/convection/diffusion into additional compartments. Moreover, from Equations (8) and (11), with $\beta_+ = 0$ we have

$$PS = \frac{\alpha_+(\alpha_+ + \alpha_-)}{\alpha_-} = K_i \qquad (14)$$

thus given the prior observation that $v_{ecs} = v_p$, with a simple relabeling, the 2 regimes generate equivalent physiological parameters in the I_3 form.

2.6 | I_2 model:
$$C^{I_2}(t) = \alpha_- e^{-t\beta_-} * C_p(t), \quad \alpha_+ = \beta_+ = 0$$

Fixing $\alpha_+ = 0$, sets $K_i = 0$; as a simplification of the AUEM, this matches the dual-input single-compartment model[19] (although now there is no *active* transfer of contrast agent). In the passive-exchange regime, the I_2 functional form is most commonly interpreted as a standard Tofts model,[18] with negligible vascular component ($v_p = 0$). However, as noted in Ref. [34], there are several other interpretations matching the same analytical form, one of which, with negligible exchange ($PS = 0$, $v_p > 0$, v_e undefined), is equivalent to the single-compartment model derived from the AUEM. Thus again, subject to a simple relabeling, the 2 regimes produce identical physiological parameters.

2.6.1 | Model optimization and selection

Given the framework outline above, there are 2 aspects to model selection: choosing a suitable analytical form, and

choosing the best interpretation of parameters given this form.

Analytical form
An analytical form can be selected examining residuals between the signal-derived and modeled concentration time series and determining which model best fits the data—choosing the model with minimum sum-of-squares errors (SSE) or applying model selection criteria such as AIC (23) to take into account model complexity.

Our results suggest that the percentage of voxels in a region attributed to each model form using AIC may itself be a useful marker. We label these as $\%_{AIC_j}$ where $j \in \{2, 3, 4 \mid I_j(t)\}$.

Physiological interpretation of parameters
We assign voxels to the active-uptake or passive-exchange regimes based on the physiologically plausible limits on compartmental volume parameters derived from the I_4 model, defining

$$I_4 - volume \quad \text{criterion} \quad \begin{cases} v_e + v_b \leq 1, & \text{assign to passive-exchange regime} \\ v_e + v_b > 1, & \text{assign to active-uptake regime} \end{cases}$$

and the percentage of all valid voxels in a region-of-interest assigned to the active-uptake regime as $\%_a$ (see Supporting Information: Selecting physiological regime for details).

2.7 | Monte Carlo simulations

Given model parameters, a set of time points and corresponding arterial and portal vein input functions (AIF/PIF), we can use Equations (5)-(13) to simulate dynamic time series. We generated multiple 1,000 sample datasets, each defined by their model form $\mathcal{F}_i$, temporal resolution δ, post-contrast duration d and noise distribution η. We used a single fixed ground truth parameter set for each of 4 model forms (units omitted for clarity):

- $\mathcal{F}_1$ AUEM (I_4): $F_p = 0.5$, $v_{ecs} = 0.1$, $K_i = 0.07$, $K_{ef} = 0.05$, $f_a = 0.2$, $\tau_a = 0.2$
- $\mathcal{F}_2$ Active-uptake only (I_3): as $\mathcal{F}_1$ except for $K_{ef} = 0$
- $\mathcal{F}_3$ Single-compartment (I_2): $F_p = 0.5$, $v_{ecs} = 0.1$, $K_i = 0$, $K_{ef} = 0$, $f_a = 0.2$, $\tau_a = 0.1$
- $\mathcal{F}_4$ 2CXM (I_4): $F_p = 0.5$, $PS = 0.1$, $v_e = 0.07$, $v_p = 0.05$, $f_a = 0.6$, $\tau_a = 0.05$

For each form, physiological parameters were converted to functional parameters ($\alpha_\pm$, $\beta_\pm$) using the active-uptake ($\mathcal{F}_1$-$\mathcal{F}_3$) and passive-exchange ($\mathcal{F}_4$) regimes. The functional parameters were then used to compute modeled concentration time series using all combinations of:

- Temporal resolution $\delta = 3.8$ seconds ($\approx$ patient dataset), 6.0 seconds ($\approx$ volunteers dataset), 12.0, 30.0 seconds.
- Post-bolus duration $d = 6$, 12, 18 minutes.

In each case, a population AIF[37] and derived PIF (Supporting Information: Input functions,[15]) were sampled at the generated times.

Each concentration time series was converted to simulated signals with baseline $S(0) = 100$ (Supporting Information: Computing signal from concentration), and replicated 1,000 times with randomly sampled Rician noise distributions $\eta_\sigma = \text{Rice}(S(t), \sigma)$ for $\sigma = 10$, 20, 30, 40. These signal data were then converted back to concentration time series.

In addition, using the group average AIF and derived PIF from the patient dataset, we generated "real" noise datasets with temporally varying noise distribution η_r, by (1) randomly sampling 1,000 voxels from across the patient dataset; (2) subtracting the fitted I_4 model concentration from the signal-derived concentration; and (3) adding the resulting residuals to the simulated concentration time series for each model form $\mathcal{F}_1$-$\mathcal{F}_4$. A second experiment using varying parameters randomly sampled from measured distributions of real data is described in Supporting Information.

2.7.1 | Model fitting

The I_2, I_3, and I_4 models were fitted to each Monte Carlo sample, and AIC used to compute $\%_{\text{AIC}_2}$, $\%_{\text{AIC}_3}$, and $\%_{\text{AIC}_4}$ for each dataset. Physiological parameter sets in the active-uptake and passive-exchange regimes were derived for each fitted model, and the $I_4 - volume$ criterion used to compute $\%_a$. In addition, for each dataset, parameter distributions were generated using (a) the AIC-selected model at each sample; and (b) the model with minimum SSE.

2.8 | In vivo data

2.8.1 | Healthy volunteers

The healthy volunteer dataset comprised of 8 volunteers (5 male, aged 18-29 years; 3 female aged 18-22 years; mean age 23 years) and were imaged with contrast injection at 2 visits on a Philips 1.5 T Achieva MRI scanner. DCE-MRI was performed (see Supporting Information: In vivo data for protocol details), with contrast agent gadoxetate disodium (Primovist/Eovist, Bayer) was administered at the 20th time point (120 seconds) using a power injector (dose 0.025 mmolkg^{-1}; 3 mLs^{-1} flow rate, flushed with 20 mL of saline at the same rate).

These data were collected for a previous study (approved by the university's ethics committee; all subjects gave informed consent prior to imaging) by Georgiou et al, with full acquisition details and parameter estimates for the active-uptake and efflux model fitted to ROI-averaged time series presented in.[31] For the experiments presented here, to enable efficient motion correction and voxel-by-voxel model fitting, the volumes were cropped to the margins of the liver, and only the first 264 time points used.

Patients with HCC

The second dataset collected for this study comprised of 10 patients with HCC (8 males and 2 females, aged 46-80 years; median 63 years) who were yet to begin treatment and had at least 1 lesion that could be measured in one dimension, according to the Response Evaluation Criteria in Solid Tumors (RECIST, see Supporting Information for eligibility criteria, tumor properties, and known co-morbidities). The study received Institutional Board Approval. Informed consent was obtained from all patients.

Each patient underwent a free-breathing coronal 3D FLASH DCE-MRI protocol on a Siemens 1.5 T Avanto system (see Supporting Information: In vivo data for protocol details). Gadoxetate was administered at the 8th time point (30 seconds) using a power injector (dose 0.025 mmolkg^{-1}; 3 mLs^{-1} flow rate).

2.8.2 | Model fitting and selection

Each dataset were motion corrected, annotated, and used to compute vascular input function (Supporting Information: Data preparation). The I_2, I_3, and I_4 models were fitted to the DCE time series of all voxels in the liver volume for each patient (Supporting Information: Model optimization). For healthy volunteers, the model framework was applied, testing 3 durations of sequence: 6 minutes post-contrast ($\approx$ patient dataset; $n_t = 80$), 12 minutes ($n_t = 138$), and 24 minutes ($n_t = 264$). Due to the computational cost of fitting models to longer time series, the models were fitted to 1000 randomly selected voxels across the liver volume, and for a spatial visualization, to all voxels in the central slice of the liver.

Model parameters were converted to physiological parameters in both regimes, and the $I_4 - volume$ criterion used to assign each voxel as active uptake or passive exchange. In addition, both AIC and minimum SSE were used to assign a model form at each voxel, with the percentage of AIC selections of each form used to analyze trends in the data, and minimum SSE used to select a single parameter set at each voxel. Parameters were summarized for each ROI taking the median over all voxels assigned to each regime.

3 | RESULTS

3.1 | Monte Carlo model selection and parameter estimation
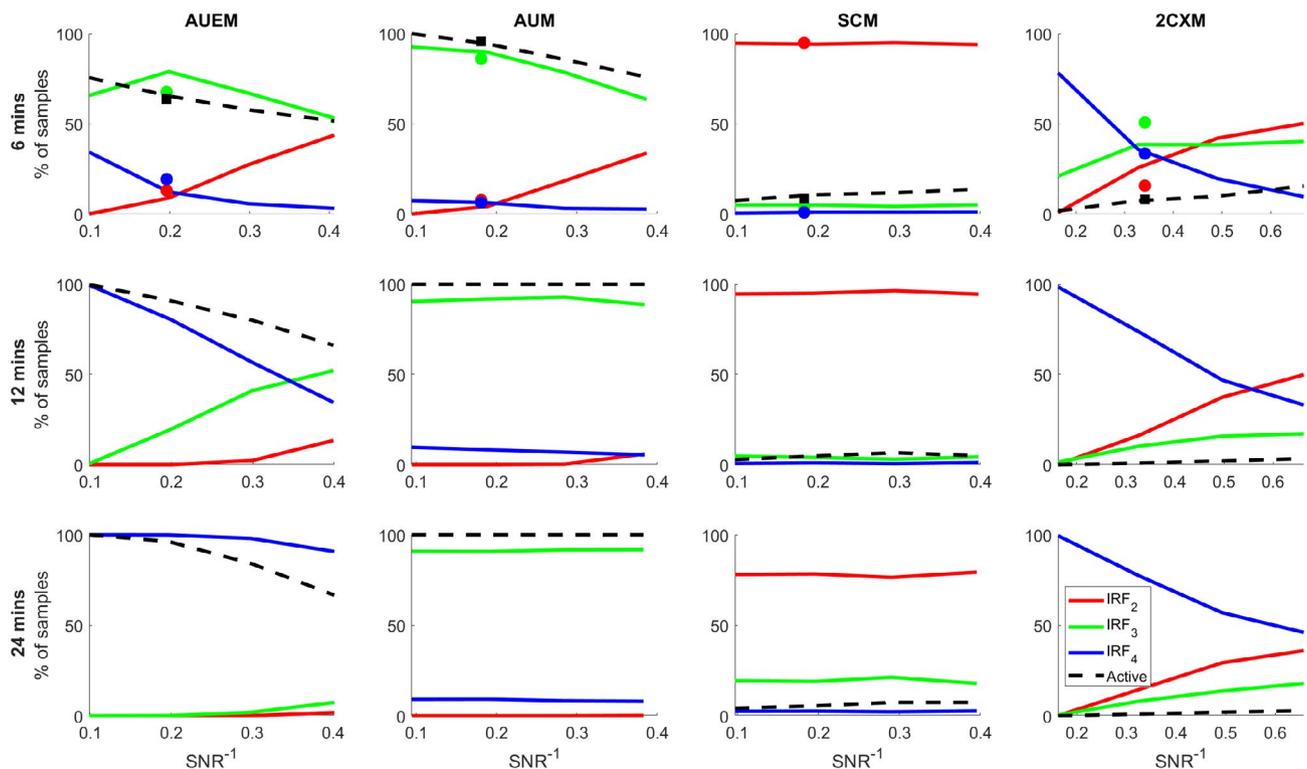
3.1.1 | Effect of temporal resolution

Increasing temporal resolution led to improved parameter estimates (Figure 3 shows K_i estimation for model form $\mathcal{F}_1$, see Supporting Information Figures S14-S37 for all other parameters), and better ability to select the ground-truth model form (Figure 2, Supporting Information Figures S7-S9). At low resolution ($\delta = 30$ seconds), with the exception of the single-compartment model form, parameter estimation is generally unreliable (at long duration, it may be possible to estimate v_{ecs} and K_i in active-uptake forms, and v_e in the passive-exchange form, although precision is significantly worse than at higher temporal resolutions). As a result, at all but the lowest noise levels, the I_2 model is selected by AIC score in the majority of samples, for all model forms (Supporting Information Figure S9).

As δ reduces from 3.8 to 12.0 seconds, some parameters appear more sensitive to the loss of resolution. For example, f_a, which is largely characterized by the shape of the

3.1.2 | Effect of duration

At short duration ($d = 6$ minutes), I_3 is the AIC-preferred functional form for both active-uptake forms at all but very low noise levels (Figure 2, Supporting Information Figure S7). Parameter predictions of F_p, v_{ecs} and K_i are consistent in both models (Supporting Information Figure S14, S17, S18). I_4 predictions of K_{ef} lack precision but are not biased (Figure S19). As the duration increases ($d = 12$ minutes), it becomes possible to resolve K_{ef}, which in turn reduces the ability of the I_3 model to properly represent the latter time points of the time series, leading to the increasing preference for the I_4 model in AIC selection (Figure 2) *and* the tendency for the I_3 model to underestimate K_i (Supporting Information Figure S18). Despite this, at noise levels observed in the real datasets, the I_3 model is still the AIC-preferred model in approximately half the samples, leading to the overall sample median of both K_i and K_{ef} being underestimates of the ground truth if AIC selection is used to generate final parameter

contrast time series in a short period after bolus injection, is affected more than parameters such as v_{ecs} and K_i (Figure 3, Supporting Information Figures S14-S19).

FIGURE 2 Analytical form and physiological regime model selection percentages for the fixed Monte Carlo datasets, with temporal resolution $\delta = 3.8$ seconds (see Supporting Information for other resolutions). Ground truth model forms are shown in columns: active uptake and efflux (AUEM); active-uptake only (AUM); single-compartment (no uptake, SCM); 2-compartment exchange (2CXM). At 6 minutes duration, solid colored circles show results for the "real" noise datasets (see main text). SNR was computed by dividing the mean signal by the root-mean-squared error of the noisy signal from the ground truth signal
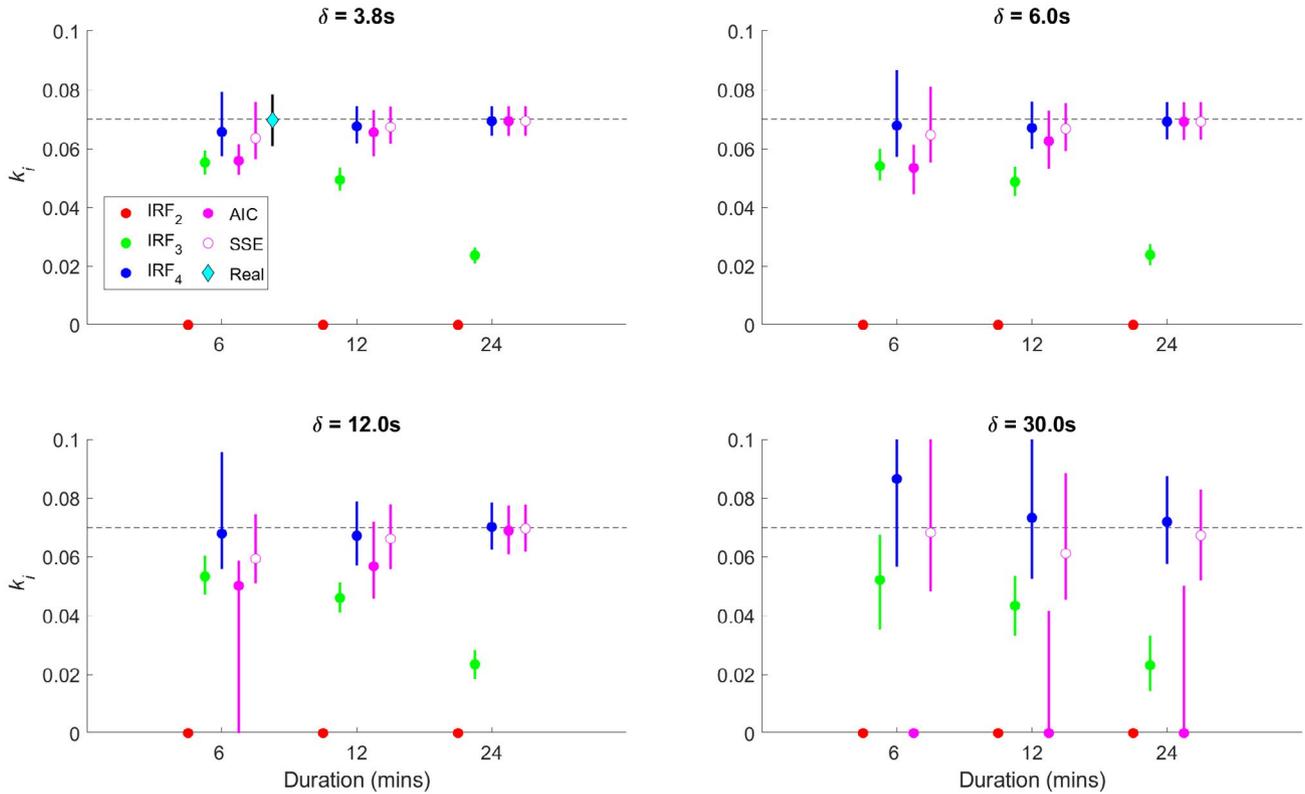
FIGURE 3 Estimation of K_i for the AUEM in the fixed Monte Carlo datasets at a single noise level (most closely matching measured real noise). Dashed black lines show parameter ground truth ($K_i = 0.07$); circles show median estimation over the 1,000 samples in each dataset using each IRF form, AIC, and minimum SSE, with vertical lines extending from the 25th to 75th percentiles. For temporal resolution $\delta = 3.8s$ and duration $d = 6$ mins, the cyan diamond and black lines show median and interquartile range for parameters estimation using minimum SSE on the dataset with directly sampled "real" noise added (see Supporting Information Figures S14-S37 for other parameters, model forms and noise levels)

estimates. This trend continues in longer durations, so that for time series $d = 24$ minutes, the I_4 is preferred in the large majority of samples (>90% at real noise levels).

Similar effects are seen for the 2-compartment exchange ground truth form. At 6 minutes, while the I_4 model is correctly selected with negligible noise, approximately half of samples at real noise levels are assigned to the I_3 model (Figure 2), producing strongly biased estimations of F_p (Supporting Information Figure S20) and (unidirectional) rate constant PS (Supporting Information Figure S23), as well as leaving v_e undefined. This is largely resolved at 12 minutes, where the I_4 model is consistently selected for the large majority of samples, and $\%_a \approx 0$ at all noise levels.

3.2 | In vivo data

3.2.1 | Healthy volunteers dataset

Model selection
Table 1 shows model selection percentages and parameter estimates for the volunteers dataset. At all durations, nearly all voxels were categorized as active-uptake (95.5% at 6 minutes, 97.8% at 12 minutes, and 98.3% at 24 minutes, Figure 4A),

thus we only include active-uptake regime parameters in further analysis.

At all durations, few voxels were best matched to the I_2 model in AIC selection (0.7%, 0.3%, and 0.2%, respectively). At 6 minutes, most voxels attributed to the I_3 model over the I_4 model (82.3% vs 16.4%), with model selection shifting to the I_4 model as duration increased (27.4% vs 71.9% at 12 minutes; 4.4% vs 94.8% at 24 minutes). These trends were spatially consistent (Figure 5A-C), and were observed in all 8 subjects, suggesting a consistent pattern of behavior in healthy liver tissue. This supports the hypothesis that at durations before efflux significantly effects gadoxetate concentration an uptake only model is sufficient to fit the data. However as duration increases the additional efflux term is necessary.[13,31]

Parameter estimates
If, as suggested by the Monte Carlo simulations, estimates for F_p, v_{ecs}, K_i, and K_{ef} are the most accurate using the I_4 model over the longest duration time series, we can observe how well estimations from the less complex models and/or shorter post-contrast durations correspond with these. I_4 estimations of K_{ef} correspond well at 12 minutes (linear correlation coefficient $\rho = 0.91$, $P < .001$,

TABLE 1 Healthy volunteer data: median (*IQR*)* values for each active-regime physiological parameter and model selection percentages

	6 minutes		12 minutes		24 minutes	
	Visit 1	**Visit 2**	**Visit 1**	**Visit 2**	**Visit 1**	**Visit 2**
Parameter						
F_p	1.06 (0.43)	0.74 (0.88)	1.01 (0.47)	0.75 (0.86)	1.01 (0.48)	0.76 (0.79)
f_a	0.17 (0.06)	0.28 (0.74)	0.17 (0.07)	0.29 (0.76)	0.19 (0.07)	0.30 (0.66)
τ_a	0.15 (0.09)	0.17 (0.02)	0.15 (0.09)	0.17 (0.02)	0.15 (0.09)	0.17 (0.02)
v_{ecs}	0.11 (0.03)	0.10 (0.02)	0.11 (0.03)	0.10 (0.02)	0.11 (0.02)	0.10 (0.02)
K_i	0.07 (0.01)	0.08 (0.03)	0.07 (0.01)	0.08 (0.03)	0.07 (0.01)	0.08 (0.02)
K_{ef}	0.00 (0.00)	0.00 (0.00)	0.02 (0.01)	0.02 (0.02)	0.02 (0.01)	0.02 (0.01)
IRF form						
$\%_{AIC_2}$	0.6 (0.6)	0.5 (0.6)	0.3 (0.4)	0.1 (0.4)	0.2 (0.4)	0.1 (0.3)
$\%_{AIC_3}$	91.3 (7.6)	89.9 (9.5)	24.3 (24.3)	23.4 (30.0)	3.5 (4.7)	2.0 (1.6)
$\%_{AIC_4}$	7.1 (7.6)	9.4 (9.6)	74.7 (23.1)	76.7 (30.4)	94.3 (4.3)	97.0 (2.0)
Regime						
$\%_a$	92.9 (4.8)	96.7 (3.1)	98.0 (3.6)	98.9 (4.0)	98.2 (1.8)	98.6 (2.9)

Notes: Units are F_p: mlmin^{-1}ml^{-1}, f_a: no units, τ_a: minute, v_{ecs}: mlml^{-1}, K_i: min^{-1}, K_{ef}: min^{-1}. In Ref. [31], Georgiou et al report mean values ($\pm s. d.$) of $F_p = 1.00 \pm 0.26$, $v_{ecs} = 0.20 \pm 0.05$, $K_i = 0.22 \pm 0.05$, $K_{ef} = 0.02 \pm 0.006$, $f_a = 0.17 \pm 0.12$. The values for v_{ecs} and K_i differ significantly from our results however this is largely due to the analysis in[31] using the blood plasma relaxivity of gadoxetate (see Discussion). *Cohort data for all parameters were tested for normality using the Anderson-Darling test. Several parameters had non-normal distributions. For consistency, median and *IQR* are used to summarize the data for all parameters.

(A) Control, 8 minutes   **(B)** Patient 7, slice 22   **(C)** Patient 9, slice 9
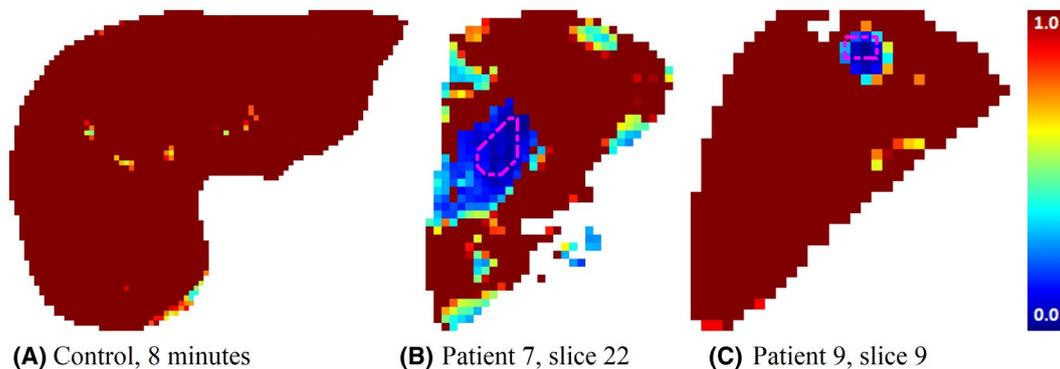
FIGURE 4 Parameter maps showing $v_e + v_b$ for a 2CXM interpretation of the I_4 functional parameters at each voxel. The colormap is capped at 1 (deep red) to highlight voxels that do not have a valid physiological interpretation in the passive-exchange regime and thus designated as active-uptake voxels. In a functioning liver, we expect the majority of voxels to be active uptake (red): this is observed in the healthy volunteer and both patients. For each patient, a single slice in the liver containing the most tumor ROI (pink dashed line) voxels is shown, with the tumor corresponding to areas where a passive-exchange regime interpretation is valid

slope $m = 1.09$, offset $c = 0.00$) but not at 6 minutes ($\rho = 0.45, P = .08, m = 0.6, c = 0.01$); I_4 estimates of F_p, v_{ecs} and K_i were accurate at both 6 and 12 minutes (all $\rho > 0.98, P < .001, m \approx 1$). I_3 estimates of F_p, v_{ecs} and K_i correspond best at 6 minutes—at 12 and 24 minutes, estimations become inaccurate because the model is insufficiently flexible to fit the plateauing concentration of gadoxetate as efflux starts to take effect, and thus the final fit reflects a compromise between fitting the early and later parts of the time series. F_p and v_{ecs} are poorly estimated by the single-compartment I_2 model at all durations.

The arterial fraction f_a proved harder to fit in some volunteer visits than either the Monte Carlo or patient data, and resulted in consistent over-estimations (ie,, $f_a \approx 1$ for the majority of voxels) in 3 DCE sequences. These data had lower SNR than the rest of the dataset, most likely due to unresolved motion artefacts. The lower temporal resolution of the volunteer data may also be a factor. Analysis suggests f_a overestimation results in F_p under-estimation; however, other parameters are largely unaffected. Parameter medians varied more between visits than between subjects, although across the group no parameter median was significantly different between visits.

(A) Healthy volunteer, 6 minutes   **(B)** 12 minutes   **(C)** 24 minutes

(D) Patient 6, slice 23   **(E)** Patient 7, slice 21   **(F)** Patient 8, slice 17
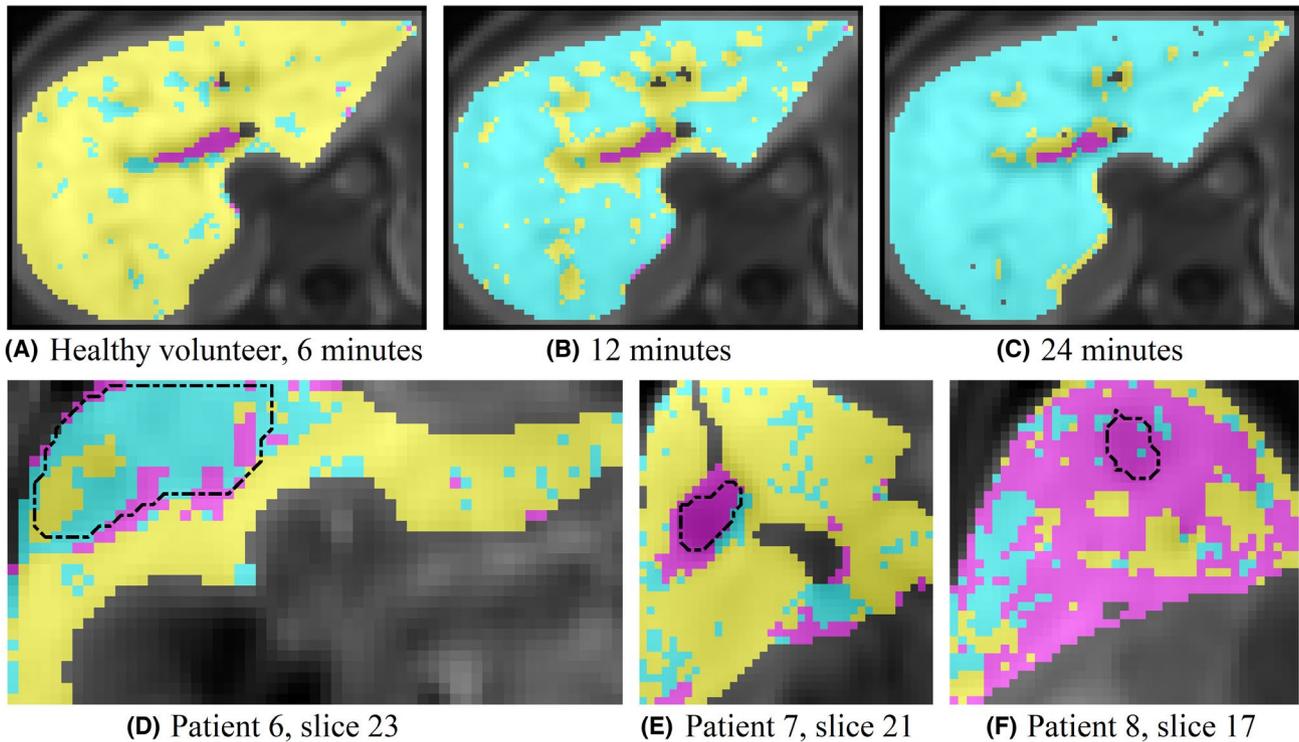
FIGURE 5   AIC selection maps for model functional form: magenta=I_2, yellow = I_3, cyan = I_4. For 1 volunteer, a the central slice in the liver is shown for post-contrast duration A, 6 minutes; B, 12 minutes; D, 24 minutes. For each patient, a single slice in the liver containing the most tumor ROI (black dashed line) voxels is shown. In normal functioning liver, we expect the majority to be I_3 (yellow). In patient 8, the preference for I_2 (magenta) over I_3 (yellow) throughout the slice highlights the loss of liver function associated with this patient's chronic liver disease. The difference in tumor composition between patient's 6 (largely I_4, cyan) and 7 (largely I_2, magenta) are also noticeable

TABLE 2   Percentage of voxels in non-tumorous liver (L, white rows) and tumor (T, shaded rows) selected by functional form ($\%_{AIC_{2,3,4}}$) and under the active-uptake interpretation ($\%_a$) for each patient

	Patient											
IRF form	Model %	1[b]	2	3	4	5[b]	6	7	8	9	10	All[c]
	$\%_{AIC_2}$ L	2.8	1.6	10.5	5.2	4.1	5.9	2.5	33.2	1.0	4.5	4.5
	T	0.0	50.0	26.7	11.2	9.1	9.9	80.2	55.2	21.9	33.9	26.7
	$\%_{AIC_3}$ L	83.9	83.9	75.1	72.1	80.3	64.3	87.4	54.2	84.9	75.6	75.6
	T	100.0	50.0	48.6	65.4	62.3	25.2	7.5	32.6	34.4	62.5	48.6
	$\%_{AIC_4}$ L	13.3	14.5	13.8	22.7	15.6	29.8	10.0	11.9	14.0	19.9	14.5
	T	0.0	0.0	24.7	11.2	26.1	55.6	12.3	12.2	43.8	3.6	12.3
Regime	$\%_a$ L	72.7	74.8	70.1	87.0	64.5	33.5	75.5	47.3	96.7	93.2	74.8
	T	100.0	53.7	25.3	38.8	31.4	3.0	4.8	5.2	28.1	55.4	28.1

[b]Tumor in patient 1 has only 6 voxels in ROI.

[b]Average of 2 tumors in patient 5.

[c]Computed pooling voxels from all patients ROIs.

3.2.2 | Patient dataset

Across the 10 patients, the I_3 model form was selected by AIC analysis in 76.2% of voxels in the (non-tumorous) liver and 48.8% of tumor voxels. The I_2 and I_4 models were more often selected in the tumor, 29.8% and 18.9%, respectively, compared with 7.1% and 16.5% in the liver. $\%_a = 71.6\%$ in the liver, compared with $\%_a = 34.6\%$ in tumors. Compared with healthy volunteers, there was more variability in individuals; Table 2 gives per patient data. Figures 4B,C and 5D,E show the spatial consistency of model selections and the extent to which they correspond with (independently demarcated) tumor borders.

TABLE 3 HCC patient data: median $(IQR)^a$ values for each physiological parameter of the active-uptake and passive-exchange regimes

Parameter	2CXM voxels		AUEM voxels		All voxels	
	Liver	Tumor	Liver	Tumor	Liver	Tumor
F_p	0.75 (0.44)	0.45 (0.19)	0.55 (0.47)	0.67 (0.60)	0.56 (0.48)	0.49 (0.46)
f_a	0.24 (0.08)	0.40 (0.81)[b]	0.17 (0.09)	0.34 (0.16)[b]	0.18 (0.09)	0.37 (0.29)[b]
τ_a	0.17 (0.06)	0.08 (0.13)[c]	0.14 (0.03)	0.10 (0.08)[c]	0.14 (0.03)	0.08 (0.12)
PS	0.07 (0.03)	0.04 (0.02)[c]	-	-	-	-
v_e	0.58 (0.17)	0.14 (0.17)[c]	-	-	-	-
v_p	0.09 (0.05)	0.07 (0.07)[c]	-	-	-	-
v_{ecs}	-	-	0.14 (0.02)	0.11 (0.03)	-	-
K_i	-	-	0.06 (0.04)	0.02 (0.02)[c]	-	-
K_{ef}	-	-	0.00 (0.00)	0.00 (0.00)	-	-

Notes: For each parameter, a per-patient value in the liver or tumor is computed taking the median of voxels in the ROI for which the model is deemed to be the correct physiological interpretation. Median values for F_p, f_a, and τ over all voxels, regardless physiological interpretation, are also shown for the liver/tumor. Units are F_p: mlmin^{-1}ml^{-1}, f_a: no units, τ_a: minute; PS: min^{-1}, v_e: mlml^{-1}, v_p: mlml^{-1}; v_{ecs}: mlml^{-1}, K_i: min^{-1}, K_{ef}: min^{-1}.

[a]Cohort data for all parameters were tested for normality using the Anderson-Darling test. Several parameters had non-normal distributions. For consistency, median and *IQR* are used to summarize the data, and paired Wilcoxon signed rank tests used to assess significance.

[b]Indicate tumor values significantly higher than the liver.

[c]Indicate tumor values significantly lower than the liver.

Parameter estimates for the group are shown in Table 3. For the parameters shared in both model interpretations (F_p, f_a, τ_a), mean plasma flow F_p was similar in liver and tumor (cohort median 0.45 vs 0.44 mlmin^{-1}ml^{-1}); arterial fraction f_a was significantly higher in tumors (0.37 vs 0.18, Willcoxon signed rank test $P = .011$), and arterial delay time significantly shorter (0.07 minutes vs 0.14 minutes, $P = .027$).

Figure 6 depicts waterfall plots of active-uptake regime physiological parameter estimates in the liver in comparison with healthy volunteers. There were no significant differences in K_i, f_a, and τ_a; however, the patient group has significantly lower F_p (Wilcoxon ranked sum test $P = .002$) and higher v_{ecs} ($P = .003$).

Pathophysiology of non-tumorous liver
In patient data, the pattern of model selection in non-tumorous tissue broadly matched the healthy volunteer cohort. However, there was a general reduction in gadoxetate uptake as seen in the average concentration time-series plots (Figure 7D,E), reflected in reduced estimates of F_p and $v_i = 1 - v_{ecs}$, higher values for $\%_{AIC_2}$ and lower values for $\%_a$. This suggests reduced perfusion and a lower volume of hepatocytes in liver tissue, which may indicate reduced function associated with co-morbidities to HCC.

This deviation from the healthy pattern was particularly strong in 2 patients, both of whom had an abnormally low $\%_a$. One of these (patient 8, $\%_a = 58.1$) also had an abnormally high proportion of I_2 voxels ($\%_{AIC_2} = 33.2$); SNR and model residuals were in the mid-range of the cohort, suggesting this was not caused by poor model fitting, but is instead consistent with restricted hepatocyte uptake of gadoxetate. This accords

with the clinical observation that patient 8 has chronic hemochromatosis, resulting in a severely cirrhotic liver.

The second (patient 6, $\%_a = 65.7$) had higher than average I_4 voxels ($\%_{AIC_4} = 29.8$), and an uptake rate in the non-tumorous liver $K_i = 0.015$ minutes^{-1} less than a third of any other patient or healthy volunteer. Again, SNR was in range, suggesting the results were a genuine reflection of abnormal liver tissue function, and is consistent with complications observed in the subsequent treatment of patient 6, in which the location and size of the tumor caused a partial obstruction of the biliary tree.

Gadoxetate uptake within HCC
While it was expected tumor voxels would be more likely than non-tumorous liver voxels to be assigned to the passive-exchange regime, there are significant variations between patients and the active-transport regime is still assigned to a significant number of tumor voxels (34.6% across the cohort). In some cases, these may correspond to misalignment in ROI boundaries, or as a result of the dynamic duration being too short to properly resolve the I_4 model form. However, an analysis of parameters for active-uptake voxels in the tumor suggests they are systematically different from those in non-tumorous tissue (Table 3). Specifically, they have increased arterial and reduced venous supply (ie, higher f_a, 0.36 vs 0.17), and significantly reduced uptake parameter K_i (0.019 vs 0.064).

Plausible interpretations of these observations are that such voxels are composed of cells exhibiting residual OATP function, allowing uptake of gadoxetate through the OATP pathway, or that there is a partial volume effect between
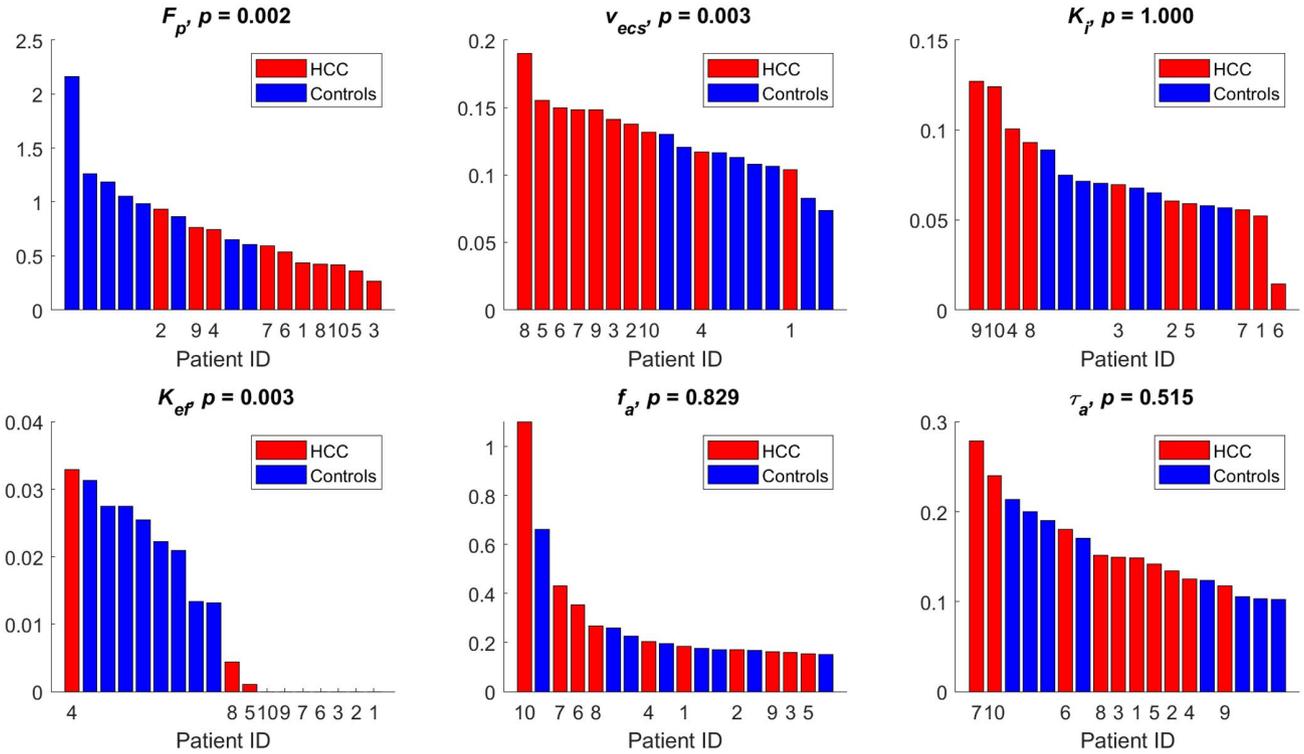
FIGURE 6 Waterfall plots of each active-uptake regime parameter for the 8 healthy volunteers and 10 HCC patients. Due to the short post-contrast duration in the patient data, estimations of K_{ef} are not reliable, and should not be compared to the volunteer data
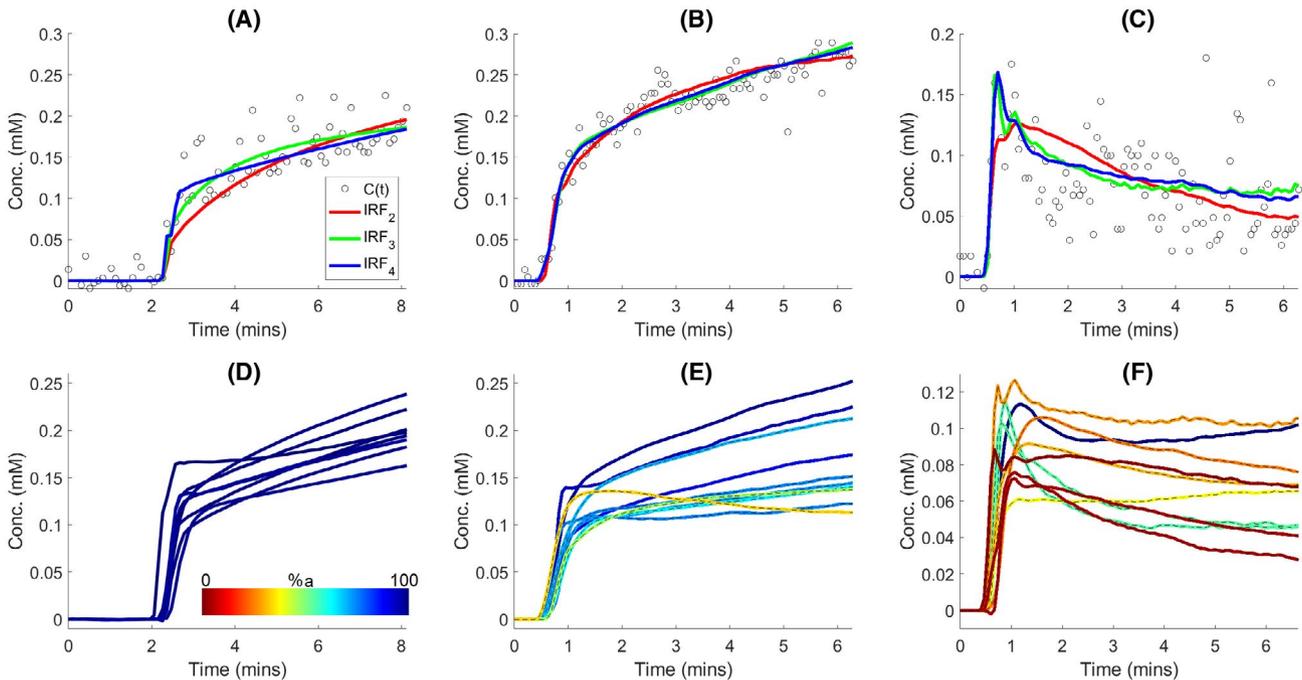
FIGURE 7 Dynamic concentration time series. Top row, sample individual voxels from A, a healthy volunteer; B, a patient liver; C, a patient HCC. Each sample show signal-derived concentration (black dots), and the fits of the I_2 (red), I_3 (green) and I_4 (blue) models. Bottom row: average modeled time series for each D, healthy volunteer liver; E, patient liver; F, patient HCC tumor. The lines are color coded with the percentage of voxels assigned to the active-uptake regime, $\%_a$. For the healthy volunteer time series in A,D only 6 minutes post-contrast is shown to compare directly with the HCC data. See Supporting information Figure S6 for the same time series shown to 24 minutes post-contrast

tumor cells and functioning liver tissue (or a combination of both). In either interpretation, if further clinical or histological information were available, a per tumor analysis of this effect may help understand clinically relevant physiological differences.

4 | DISCUSSION

Analysis of the in vivo data conforms to our initial hypotheses and reflects the results of the Monte Carlo simulations: dynamic time series in healthy liver tissue are best fitted by models that account for active transport of gadoxetate by hepatocytes from a well-mixed arterial and venous blood supply. In comparison, tumors tend to exhibit passive-exchange of contrast agent, with a more rapid and arterially dominated blood supply. However, there are also significant differences within the non-tumorous liver—where chronic disease such as cirrhosis and fibrosis may reduce (or even eliminate) portal blood supply or inhibit hepatoctye function—and within tumor sub-types.

These results highlight the necessity of fitting multiple models to characterize tissue across the liver volume. Relying on a single preselected model interpretation risks fitting a model that does not adequately match the data and thus produces unreliable parameter estimates. Instead, fitting models in the functional space, and then post-fit, choosing the most appropriate physiological interpretation at each voxel provides a more robust analysis, capable of dealing with variations between individual patients and between/within heterogeneous tissues. Moreover, the data, and in particular the varying ways in which patients 6 and 8 differ from the cohort norm, suggest analyzing both the appropriate physiological interpretation *and* the selected level of model complexity provides clinically relevant information.

Analyzing model functional form helps clarify the relationship between different physiological interpretations and their associated assumptions of tissue microstructure. From a practical perspective, it is much more efficient to extract multiple physiological derivations from a single optimization, as the former has negligible cost compared with the computationally intensive process of fitting a complex non-linear model to each voxel. The functional forms in our framework do not allow direct measurement of sinusoidal backflux. This could be incorporated by extending the top-level 2-compartment models to a 3-compartment model,[13] although this would significantly increase computation and may require longer acquisitions to provide sufficient data points to unambiguously resolve all model parameters.

In common with many previous studies of DCE-MRI in the liver, we did not take steps to suppress liver fat in these data. This may lead to errors in T_1 estimation and that may subsequently bias the model parameters. The relativity of gadoxetate should also be carefully considered. This is reported as 6.9 mmol^{-1} s^{-1} plasma at 1.5 T[38]; however, a subsequent study measured values approximately twice as high in liver tissue (14.0 mmol^{-1} s^{-1},[39]). We used the former for converting signal to concentration for AIF generation, and the latter for model fitting in the liver (see Supporting Information), while accepting the limitation that patient disease may also effect gadoxetate relaxivity (eg, due to reduced albumin).

With only 10 patients imaged once using a 6-minute post-contrast acquisition, there are limits to the conclusions that can be drawn from the HCC data. A larger cohort, longer DCE protocol, with repeat baseline imaging would enable us to assess the repeatability of the analysis, while follow-on imaging would allow testing of the clinical utility of specific parameters as biomarkers in tracking disease progression or treatment response. However, our results correlate with the clinical presentation of the patients. Moreover, the spatial consistency of model selection and parameter maps suggests that the framework is suitable for quantifying the microvascular characteristics of both non-tumorous liver tissue and HCCs on a per voxel basis, despite the challenging data quality resulting from a free-breathing protocol. Thus we propose that while HCC patients are a study group of clinical significance, there is sufficient evidence to repeat the analysis described here in a larger cohort featuring longitudinal imaging.

4.1 | Recommendations for future studies

The in vivo data, together with the Monte Carlo simulations, suggest using a DCE protocol of at least 12 minutes post-contrast when using gadoxetate as a contrast agent. This allows estimation of efflux (K_{ef}) in the active-uptake regime, and significantly improved estimation of the extracellular, extravascular volume fraction v_e in the 2CXM, as well as improving the differentiation between the 2 regimes in potentially ambiguous tissue (eg,, in HCC, or compromised non-tumorous liver).

If only shorter acquisitions are possible, as with the patient data in this study, it must be accepted that estimations of K_{ef} will lack precision, and there will be greater ambiguity in determining the appropriate physiological regime, and thus care should be taken in any conclusions drawn from the data. Where longer protocols are used, it is necessary to fit a full biexponential (I_4) model form (or alternative model with sufficient complexity) to describe both the initial uptake and subsequent plateauing of gadoxetate concentration. A simpler mono-exponential (I_3) form may appear to fit the time series (and may even by preferred under AIC selection in a significant proportion of voxels), but parameter estimates derived from the I_3 model will likely be biased.

A corollary of these observations is that if using a nested model approach, it is better to use the model with minimum SSE (rather than model selection criteria such as AIC or an F-test) from which to estimate physiological parameters. Despite this, there are still potential advantages in fitting the nested I_3 and I_2 models and applying AIC selection. A consistent pattern in model selection was observed in non-tumorous liver tissue across the healthy volunteers and in patients that did not have clinical observations of significant liver co-morbidity. This suggests that deviations from this pattern are a sign that either there is insufficient data quality to fit a full I_4 model to individual voxels, or that there may be clinically relevant reduction in liver function to assess. Such observations may be noticed more easily by a simple check on the model selection percentages than trying to interpret a more complicated distribution of multi-variate physiological parameters. Although there is a computational cost to fitting the additional models, optimization search space scales combinatorially with additional parameters, and thus the I_2 and I_3 models do not add significant overhead to fitting the I_4 model.

As with the limitations of AIC analysis, it is not possible to state *definitively* the most appropriate physiological interpretation at all voxels, and it is inevitable some will be miscategorized by the framework. Nevertheless we propose it is beneficial to compute summary statistics over only those voxels assigned to each regime, and present these with the regime's percentage composition. The alternative, preselecting a model for each region, and then averaging over all voxels, will almost certainly include more voxels that do not meet the physiological assumptions of the model, with greater risk of producing invalid parameter summary statistics.

The difficulty in fitting f_a to individual voxels in some healthy volunteer datasets, together with the results of varying temporal resolution in the fixed Monte Carlo simulations (Figure 3), suggest a temporal resolution of at least 6 seconds (and ideally higher) is necessary to adequately separate the arterial and venous components of vascular input to the liver. If accurate estimations of f_a (and τ_a) are not required, a temporal resolution of 12 seconds appears sufficient to estimate all other parameters in the framework (above 12 seconds and estimations of several other parameters become unreliable).

5 | CONCLUSIONS

Although the individual tracer kinetic models used in this study are not new, we believe their presentation here—grouped by functional form, cross-matched with vascular input and then physiologically interpreted post-fitting on a per-voxel basis—is novel and reveals a general framework within which to understand the family of models. Extensive Monte Carlo simulations, and the application of the framework to real datasets show the value in this approach, and that attempting to model the dynamics of gadoxetate in HCC using a single model, or with fixed prior assumptions about which model will be appropriate in a specific region, may not adequately cope with the varying physiology of individual patients, and ultimately will produce unreliable tracer-kinetic parameter estimates.

The unique physiology of HCC, together with varying levels of associated chronic disease in the surrounding liver, makes these tumors a challenging target for quantitative DCE imaging. We have shown that using gadoxetate as a contrast agent and applying a framework of tracer kinetic models, it is possible to investigate and quantify different aspects of tumor microvascular and general liver function in a free-breathing DCE protocol tolerated by patients. However, careful consideration should be given to sequence length and temporal resolution, to maximize the ability of model selection to differentiate between potentially ambiguous tissue types and increase the robustness of parameter estimation.

CONFLICT OF INTEREST
G.J.M. Parker has a shareholding and part time appointment and directorship at Bioxydyn Ltd. which provides MRI services and has a commercial interest in MRI-based biomarkers.

DATA AVAILABILITY STATEMENT
The code that supports the findings of this study isopenly available at https://gitlab.com/manchester_qbi/manchester_qbi_public/gadoxetate_model_selection_framework.This code uses Madym to perform DCE-MRI analysis, an open-source toolbox available at https://gitlab.com/manchester_qbi/manchester_qbi_public/madym_cxx.

ORCID
Michael Berks http://orcid.org/0000-0003-4727-2006

TWITTER

Michael Berks 🐦 @michaelberks

REFERENCES

1. Chanyaputhipong J, Low SCA, Chow PKH. Gadoxetate acid-enhanced MR imaging for HCC: a review for clinicians. *Int J Hepatol*. 2011;2011:489342.

2. Huppertz A, Haraida S, Kraus A, et al. Enhancement of focal liver lesions at gadoxetic acid-enhanced MR imaging: correlation with histopathologic findings and spiral CT-initial observations. *Radiology*. 2005;234:468-478.

3. Zech CJ, Herrmann KA, Reiser MF, Schoenberg SO. MR imaging in patients with suspected liver metastases: value of liver-specific contrast agent Gd-EOB-DTPA. *Magn Reson Med Sci*. 2007;6:52.

4. Tsuboyama T, Onishi H, Kim T, et al. Hepatocellular carcinoma: hepatocyte-selective enhancement at gadoxetic acid-enhanced MR imaging-correlation with expression of sinusoidal and canalicular transporters and bile accumulation. *Radiology*. 2010;255:824-833.

5. McInnes M, Hope TA, Vu KN, et al. Magnetic resonance imaging performed with gadoxetate disodium for the diagnosis of hepatocellular carcinoma in cirrhotic and non-cirrhotic patients. *Cochrane Database Systematic Rev*. 2017;2017.

6. Mittal S, El-Serag HB. Epidemiology of HCC: consider the population. *J Clin Gastroenterol*. 2013;47:S2.

7. Hanna RF, Miloushev VZ, Tang A, et al. Comparative 13-year meta-analysis of the sensitivity and positive predictive value of ultrasound, CT, and MRI for detecting hepatocellular carcinoma. *Abdominal Radiol*. 2016;41:71-90.

8. Li J, Wang J, Lei L, Yuan G, He S. The diagnostic performance of gadoxetic acid disodium-enhanced magnetic resonance imaging and contrast-enhanced multi-detector computed tomography in detecting hepatocellular carcinoma: a meta-analysis of eight prospective studies. *Eur Radiol*. 2019;29:6519-6528.

9. Kumada T, Toyoda H, Tada T, et al. Evolution of hypointense hepatocellular nodules observed only in the hepatobiliary phase of gadoxetate disodium-enhanced MRI. *Am J Roentgenol*. 2011;197:58-63.

10. Kim KA, Kim MJ, Jeon HM, et al. Prediction of microvascular invasion of hepatocellular carcinoma: usefulness of peritumoral hypointensity seen on gadoxetate disodium-enhanced hepatobiliary phase images. *J Magn Reson Imaging*. 2012;35:629-634.

11. Choi JY, Kim MJ, Park YN, Lee JM, Yoo SK, Rha SY, et al. Gadoxetate disodium-enhanced hepatobiliary phase MRI of hepatocellular carcinoma: correlation with histological characteristics. *Am J Roentgenol*. 2011;197:399-405.

12. Asenbaum U, Ba-Ssalamah A, Mandorfer M, et al. Effects of Portal Hypertension on Gadoxetic Acid-Enhanced Liver Magnetic Resonance: Diagnostic and Prognostic Implications. *Investigative Radiol*. 2017;52:462-469.

13. Leporq B, Daire JL, Pastor CM, et al. Quantification of hepatic perfusion and hepatocyte function with dynamic gadoxetic acid-enhanced MRI in patients with chronic liver disease. *Clin Sci*. 2018;132:813-824.

14. O'Connor J, Rose C, Jackson A, et al. DCE-MRI biomarkers of tumour heterogeneity predict CRC liver metastasis shrinkage following bevacizumab and FOLFOX-6. *British J Cancer*. 2011;105:139.

15. Banerji A, Naish JH, Watson Y, Jayson GC, Buonaccorsi GA, Parker GJM. DCE-MRI model selection for investigating disruption of microvascular function in livers with metastatic disease. *J Magn Reson Imaging*. 2012;35:196-203.

16. Jayson GC, Zhou C, Backen A, et al. Plasma Tie2 is a tumor vascular response biomarker for VEGF inhibitors in metastatic colorectal cancer. *Nat Commun*. 2018;9:4672.

17. O'Connor JP, Jackson A, Parker GJ, Roberts C, Jayson GC. Dynamic contrast-enhanced MRI in clinical trials of antivascular therapies. *Nat Reviews Clin Oncol*. 2012;9:167.

18. Tofts PS. Modeling tracer kinetics in dynamic Gd-DTPA MR imaging. *J Magn Reson Imaging*. 1997;7:91-101.

19. Materne R, Van Beers BE, Smith AM, et al. Non-invasive quantification of liver perfusion with dynamic computed tomography and a dual-input one-compartmental model. *Clin Sci*. 2000;99:517-525.

20. Patel J, Sigmund EE, Rusinek H, Oei M, Babb JS, Taouli B. Diagnosis of cirrhosis with intravoxel incoherent motion diffusion MRI and dynamic contrast-enhanced MRI alone and in combination: Preliminary experience. *J Magn Reson Imaging*. 2010;31:589-600.

21. Koh TS, Thng CH, Lee PS, et al. Hepatic metastases: in vivo assessment of perfusion parameters at dynamic contrast-enhanced MR imaging with dual-input two-compartment tracer kinetics model. *Radiology*. 2008;249:307-320.

22. Koh TS, Thng CH, Hartono S, et al. Dynamic contrast-enhanced MRI of neuroendocrine hepatic metastases: a feasibility study using a dual-input two-compartment model. *Magn Reson Med*. 2011;65:250-260.

23. Akaike H. A new look at the statistical model identification. *IEEE Trans Automatic Control*. 1974;19:716-723.

24. Blumgart LH, D'Angelica M, DeMatteo R, et al. Regional chemotherapy for unresectable primary liver cancer: results of a phase II clinical trial and assessment of DCE-MRI as a biomarker of survival. *Ann Oncol*. 2009;20:1589-1595.

25. Taouli B, Johnson RS, Hajdu CH, et al. Hepatocellular carcinoma: perfusion quantification with dynamic contrast-enhanced MRI. *Am J Roentgenol*. 2013;201:795-800.

26. Chen J, Chen C, Xia C, et al. Quantitative free-breathing dynamic contrast-enhanced MRI in hepatocellular carcinoma using gadoxetic acid: correlations with Ki67 proliferation status, histological grades, and microvascular density. *Abdominal Radiol*. 2018;43:1393-1403.

27. Nilsson H, Nordell A, Vargas R, Douglas L, Jonas E, Blomqvist L. Assessment of hepatic extraction fraction and input relative blood flow using dynamic hepatocyte-specific contrast-enhanced MRI. *J Magn Reson Imaging*. 2009;29:1323-1331.

28. Nilsson H, Blomqvist L, Douglas L, Nordell A, Jonas E. Assessment of liver function in primary biliary cirrhosis using Gd-EOB-DTPA-enhanced liver MRI. *HPB*. 2010;12:567-576.

29. Sourbron S, Sommer WH, Reiser MF, Zech CJ. Combined quantification of liver perfusion and function with dynamic gadoxetic acid-enhanced MR imaging. *Radiology*. 2012;263:874-883.

30. Sommer WH, Sourbron S, Huppertz A, Ingrisch M, Reiser MF, Zech CJ. Contrast agents as a biological marker in magnetic resonance imaging of the liver: conventional and new approaches. *Abdominal Imaging*. 2012;37:164-179.

31. Georgiou L, Penny J, Nicholls G, et al. Quantitative assessment of liver function using gadoxetate-enhanced magnetic resonance imaging: monitoring transporter-mediated processes in healthy volunteers. *Invest Radiol*. 2017;52:111-119.

32. Luypaert R, Ingrisch M, Sourbron S, de Mey J. The Akaike information criterion in DCE-MRI: Does it improve the haemodynamic parameter estimates? *Phys Med Biol*. 2012;57:3609-3628.

33. Sourbron S, Ingrisch M, Siefert A, Reiser M, Herrmann K. Quantification of cerebral blood flow, cerebral blood volume, and blood-brain-barrier leakage with DCE-MRI. *Magn Reson Med*. 2009;62:205-217.

34. Sourbron SP, Buckley DL. On the scope and interpretation of the Tofts models for DCE-MRI. *Magn Reson Med*. 2011;66:735-745.

35. Mortimore GE, Hutson NJ, Surmacz CA. Quantitative correlation between proteolysis and macro- and microautophagy in mouse hepatocytes during starvation and refeeding. *Proc Nat Acad Sci*. 1983;80:2179-2183.

36. Watanabe N, Tsukada N, Smith CR, Phillips MJ. Motility of bile canaliculi in the living animal: implications for bile flow. *J Cell Biol*. 1991;113:1069-1080.

37. Parker GJ, Roberts C, Macdonald A, Buonaccorsi GA, Cheung S, Buckley DL, et al. Experimentally-derived functional form for a population-averaged high-temporal-resolution arterial input function for dynamic contrast-enhanced MRI. *Magn Reson Med*. 2006;56:993-1000.

38. Rohrer M, Bauer H, Mintorovitch J, Requardt M, Weinmann HJ. Comparison of magnetic properties of MRI contrast media solutions at different magnetic field strengths. *Investigative Radiol*. 2005;40:715-724.

39. Jost G, Schuetz G, Pietsch H. Gadoxetate relaxivities increase significantly after hepatic uptake at clinical field strength impacting kinetic modelling for liver function analysis. In: Proceedings of the 27th Annual Meeting of ISMRM Paris, France. 2018:5010.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the Supporting Information section.

FIGURE S1 A, Liver region in central slice of a dynamic volume acquired approximately 1 minute post-contrast. B,C Per voxel estimates of temporal-signal error

FIGURE S2 Mean translation along each axis of the deformation field predicted to register each DCE volume to the target volume

FIGURE S3 Measuring a patient-specific AIF

FIGURE S4 Comparing measured vs derived PIFs

FIGURE S5 Residual errors for estimating temporal noise, before and after motion correction

FIGURE S6 Dynamic concentration time series for healthy volunteer data at 24 minutes post-contrast

FIGURE S7-S9 Analytical form (by AIC) and physiological regime selection percentages for the fixed Monte Carlo datasets, at temporal resolutions $\delta = \{6.0s, 12.0s, 30.0s\}$

FIGURE S7-S13 Analytical form (by minimum SSE) and physiological regime selection percentages for the fixed Monte Carlo datasets, with temporal resolution $\delta = \{3.8s, 6.0s, 12.0s, 30.0s\}$

FIGURE S7-S37 Estimation of active-uptake and passive-exchange regime parameters for the fixed Monte Carlo ground-truth forms $\mathcal{F}_1$ - $\mathcal{F}_4$

FIGURE S7-S49 Error box plots for varying ground truth active-uptake and passive-exchange regime parameters (5000 samples), at 6-, 12-, and 18-minute durations