Medical Image Analysis 107 (2026) 103768

Contents lists available at ScienceDirect

Medical Image Analysis

journal homepage: www.elsevier.com/locate/media

Acquisition-independent deep learning for quantitative MRI parameter estimation using neural controlled differential equations

Daan Kuppens [a,b,*], Sebastiano Barbieri [c,d], Daisy van den Berg [a,e], Pepijn Schouten [a], Harriet C. Thoeny [f,g], Hanneke W.M. van Laarhoven [h,i], Myrte Wennen [a], Oliver J. Gurney-Champion [a,b]

[a] Department of Radiology and Nuclear Medicine, Amsterdam University Medical Center, Amsterdam, The Netherlands
[b] Imaging and Biomarkers, Cancer Center Amsterdam, Amsterdam, The Netherlands
[c] Queensland Digital Health Centre, University of Queensland, Brisbane, Australia
[d] Centre for Big Data Research in Health, UNSW Sydney, Sydney, Australia
[e] Department of Biomedical Engineering and Physics, Amsterdam University Medical Center, Amsterdam, The Netherlands
[f] University Teaching and Research Hospital, University of Fribourg, Fribourg, Switzerland
[g] Department of Urology, Inselspital, University of Bern, Switzerland
[h] Department of Medical Oncology, Amsterdam University Medical Center, Amsterdam, The Netherlands
[i] Cancer Center Amsterdam, Amsterdam, The Netherlands

ARTICLE INFO

Keywords:
Quantitative MRI
Parameter estimation
Deep learning
Neural controlled differential equations

ABSTRACT

Deep learning has proven to be a suitable alternative to least squares (LSQ) fitting for parameter estimation in various quantitative MRI (QMRI) models. However, current deep learning implementations are not robust to changes in MR acquisition protocols. In practice, QMRI acquisition protocols differ substantially between different studies and clinical settings. The lack of generalizability and adoptability of current deep learning approaches for QMRI parameter estimation impedes the implementation of these algorithms in clinical trials and clinical practice. Neural Controlled Differential Equations (NCDEs) allow for the sampling of incomplete and irregularly sampled data with variable length, making them ideal for use in QMRI parameter estimation. In this study, we show that NCDEs can function as a generic tool for the accurate estimation of QMRI parameters, regardless of QMRI sequence length, configuration of independent variables and QMRI forward model (variable flip angle *T1*-mapping, intravoxel incoherent motion MRI, dynamic contrast-enhanced MRI). NCDEs achieved lower mean squared error than LSQ fitting in low-SNR simulations and in vivo in challenging anatomical regions like the abdomen and leg, but this improvement was no longer evident at high SNR. When NCDEs improve parameter estimation, they tend to do so by reducing the variance in estimation errors. These findings suggest that NCDEs offer a robust approach for reliable QMRI parameter estimation, especially in scenarios with high uncertainty or low image quality. We believe that with NCDEs, we have solved one of the main challenges for using deep learning for QMRI parameter estimation in a broader clinical and research setting.

1. Introduction

Imaging biomarkers, such as quantitative MRI (QMRI) parameters, offer accessible, cost-effective, reproducible and non-invasive tools to assist in lesion detection, characterization and treatment monitoring of various pathologies, improving patient care. QMRI techniques produce parameters used to assess tissue morphology, biology, and function. Proven useful clinical applications include oncological monitoring (Chauvie et al., 2023; O'Connor et al., 2017) and the imaging of stroke (Albers, 1998), detection of myocardial abnormalities (Manfrini et al., 2021; Taylor et al., 2016) and iron overload (Liden et al., 2021). Commonly studied QMRI techniques include the intra-voxel incoherent motion (IVIM MRI) model for diffusion-weighted imaging (DWI), *T1*-relaxometry using variable flip angle (VFA *T1*-mapping) and dynamic contrast-enhanced MRI (DCE-MRI).

In QMRI, tissue properties are estimated from a series of MRI data using biophysical models that relate the measured MRI signal to the underlying tissue properties via QMRI parameters. Conventionally, such

* Corresponding author.
E-mail address: d.kuppens@amsterdamumc.nl (D. Kuppens).

https://doi.org/10.1016/j.media.2025.103768
Received 16 December 2024; Received in revised form 7 August 2025; Accepted 10 August 2025
Available online 11 August 2025

parameters are estimated with least squares fitting (LSQ) to retrieve the QMRI parameters from MR images with different contrast weightings. LSQ fitting is an iterative process that minimizes the sum of squared differences between observed MRI data and the reconstructed signal curve. Although LSQ fitting is a reliable estimator of physiological parameters when SNR is high, it does have significant limitations. When SNR is low, the combined effect of ignoring spatial information, noisy signal curves and non-convex objective functions leads to high variance in the QMRI parameter estimates (Barbieri et al., 2016; Neil and Bretthorst, 1993; While, 2017). The compromised repeatability and accuracy of parameter estimates form a crucial hurdle for the clinical application of QMRI techniques (Kurland et al., 2012; Rosenkrantz et al., 2015).

Recent work demonstrated that deep learning, with its capacity to learn nonlinear mappings, is a suitable alternative to LSQ fitting for estimating parameters in many QMRI models. In particular, it improves accuracy and precision, yields faster parameter estimation and reduces day-to-day variation in patients (Barbieri et al., 2020; Bliesener et al., 2020; Gurney-Champion et al., 2022; Kaandorp et al., 2021; Ottens et al., 2022; Ulas et al., 2019). However, in contrast to LSQ fitting these deep learning implementations are not robust to different MR acquisition protocols, as they are dependent on the input being either a fixed set of input signals (for e.g. fully connected networks (Bliesener et al., 2020; Kaandorp et al., 2021; Ulas et al., 2019) and convolutional networks (Huang, 2022; Vasylechko et al., 2022)) or a series of regularly sampled signals (for recurrent neural networks (Ottens et al., 2022)). In practice, QMRI acquisition protocols differ substantially between studies and clinical settings (Ljimani et al., 2020), and acquisition settings do not typically follow equidistant sampling patterns (e.g. flip angles 2, 5, 10, 25°). Current deep learning approaches can therefore be considered acquisition-specific, meaning that although model architecture can be re-used, each acquisition protocol requires a dedicated model, with dedicated training data and corresponding training and validation phases. The lack of generalizability and adoptability of current deep learning approaches for QMRI parameter estimation impedes the implementation of these algorithms in clinical trials and clinical practice. Particularly, the need for acquisition-specific models requires individual training and validation phases each time scanner settings are changed, which not only adds complexity but also increases the likelihood of inconsistent results across different protocols. Hence, an acquisition-independent approach is crucial for implementation of deep learning for QMRI parameter estimation in the clinical workflow.

In parallel, a group of Neural Ordinary Differential Equations has been developed as machine learning methods that approximate system dynamics in continuous time by training a neural network to learn the underlying differential equation (Chen, 2018). Neural Controlled Differential Equations (NCDEs) build on this framework by incorporating incoming data to control the learnt differential equation with observations and thereby creating an explicit dependence of the output on the learnt system dynamics and the input series (Kidger, 2020). NCDEs are well-suited to handle irregular, incomplete, and variable-length data by using cubic splines to transform discrete inputs into a continuous, smooth trajectory, enabling the modeling of hidden state evolution through controlled differential equations. This capability makes NCDEs a promising approach for QMRI parameter estimation.

We hypothesize that NCDEs can function as generic acquisition-independent networks for QMRI parameter estimation. This would solve the abovementioned shortcomings and pave the way for the integration of deep learning for QMRI in the clinical workflow.

Our main contributions are:

- We overcome the limitation that neural networks are specific to MR acquisition protocols by implementing NCDEs for QMRI parameter estimation.
- We demonstrate NCDE performance on simulated data for VFA *T1*-mapping, IVIM MRI and extended Tofts-Kety DCE-MRI.
- We demonstrate NCDE performance on in vivo data for VFA *T1*-mapping and IVIM MRI.
- The experimental results demonstrate the superiority of NCDEs compared to conventional LSQ fitting for QMRI parameter estimation, both in simulation and in vivo.
- We demonstrate that NCDEs have similar performance to acquisition-specific multi-layer perceptrons for QMRI parameter estimation, both in simulation and in vivo.

2. Methods

All analyses were performed using Python (v3.8, Python Software Foundation) and PyTorch (v1.13.1, PyTorch Foundation). Our code is available at: https://github.com/DKuppens/NCDE-QMRI.

2.1. Quantitative MRI

QMRI models allow us to assess tissue properties by using biophysical models to describe the MRI signal intensity as a function of a changing independent variable, such as flip angle (FA) in VFA *T1*-mapping, diffusion weighting (b-value) in IVIM MRI and time (t) in DCE-MRI.

In VFA *T1*-mapping, the longitudinal spin relaxation time ($T1$) is estimated by acquiring multiple spoiled gradient-echo readouts, each with different excitation flip angles (FA). Consequently, the signal (s) at FA depends on $T1$, the repetition time (TR), and the magnetization at thermal equilibrium (s_0) (Christensen, 1974; Gupta, 1977):

$$s(FA,\ TR) = s_0 \frac{1-\exp\left(-\frac{TR}{T1}\right)}{1-\cos(FA)\ \exp\left(-\frac{TR}{T1}\right)} \sin(FA) \tag{1}$$

In IVIM MRI, diffusion-weighted gradients are used to sensitize the MRI signal to incoherent motion. By varying the strength and duration of the diffusion-weighted gradients, we are able to separate the effect of incoherent motion from diffusion with a bi-exponential model where the signal (s) intensity at diffusion weighting (b) depends on the diffusion coefficient (D), which indicates tissue cell density and structural integrity, the pseudo-diffusion coefficient (D^*) and perfusion fraction (f), which reflect tissue vascularization and perfusion, and the baseline signal intensity (s_0):

$$s(b) = s_0((1-f)\ \exp(-bD) + f\ \exp(-bD^*)). \tag{2}$$

In DCE-MRI, tissue $T1$ is measured over time after bolus administration of a contrast agent, gadolinium. Gadolinium shortens the tissue $T1$, allowing the tissue gadolinium concentration over time to be estimated. In turn, the pharmacokinetics of gadolinium reflect the tissue perfusion and blood vessel permeability (Khalifa et al., 2014; Sourbron and Buckley, 2013). In the extended Tofts-Kety Model for DCE-MRI, tissue gadolinium concentration (C_t) at time t depends on the reflux rate (k_e), fractional volume of extravascular extracellular space (v_e), fractional volume of plasma (v_p) and the concentration of gadolinium in blood plasma (C_p) (Kety, 1951; Tofts, 1999):

$$C_t(t, C_p) = v_p C_p(t) + k_e v_e \int_0^t C_p(\tau)\ \exp(-k_e(t-\tau))d\tau. \tag{3}$$

For consistency between QMRI forward models throughout the text, we will refer to C_t as s. For the extended Tofts-Kety DCE-MRI model, we utilized the implementation and fitting routines from the Python package OG_MO_AUMC_ICR_RMH_NL_UK (Orton et al., 2008), available on the OSIPI GitHub repository (van Houdt et al., 2023). In all described experiments, C_p remained fixed and was equal to the version used in (Rata et al., 2015) (for detailed information, see Table S4).

2.2. Evaluation methods

2.2.1. NCDE model

The proposed approach assumes the existence of a hidden space where changes in the measured signal curve can be described by a controlled differential equation. NCDEs are able to estimate QMRI parameters independent of the acquisition scheme by learning to parameterize the differential equation that governs the evolution in hidden space of the system responsible for generating the observed measurements. More specifically, NCDEs (Fig. 1) learn a parameterized mapping from the input sequence S to output y, using three fully connected neural networks ($l^1_\varnothing$, $f_\varnothing$ and $l^2_\varnothing$). The input sequence (S) contains the signal intensities (s, either a series of MRI signals for VFA *T1*-mapping and IVIM MRI or a derived gadolinium concentration curve for DCE-MRI) along with corresponding values of the independent variable (v, either the flip angle for VFA *T1*-mapping, b-value for IVIM MRI or time for DCE-MRI). S was organized in ascending order based on the values of the independent variable. The input sequence defines the knots of a twice continuously differentiable cubic spline (X), representing an approximation of the continuous underlying process which is observed through S. Although fixed sampling of the cubic spline could provide standardized input lengths, poor extrapolation combined with variability in ranges of the independent variable would cause inconsistencies in sample accuracy, interval and frequency. NCDEs are therefore required to robustly model the signal across varying acquisition protocols. $l^1_\varnothing$ maps the first element of the input sequence (X_0) to the initial value in hidden space (z_{t_0}). An auxiliary variable (τ) is introduced to function as a placeholder for the independent variable. The hidden state (z) then evolves within the hidden space according to the product of $f_\varnothing(z)$ and the derivative of the input sequence with respect to the auxiliary variable ($\frac{dX}{d\tau}$). $f_\varnothing$ represents a learned vector field in hidden space. The hidden state at t (z_t) is then obtained by solving the controlled $f_\varnothing$ over the distance between t_0 and t using an adaptive step size solver, following:

$$z_t = z_{t_0} + \int_{t_0}^{t} f_\varnothing(z(\tau))\frac{dX}{d\tau}(\tau)d\tau. \tag{4}$$

$l^2_\varnothing$ reads out the final hidden state and maps it to an output vector (y) containing coefficients of the estimated QMRI parameters.

In our implementation, we kept the architecture similar to the original implementation of NCDEs (Kidger, 2020). We made use of the *NeuralCDE* (https://github.com/patrick-kidger/NeuralCDE) and *torchcde* (https://github.com/patrick-kidger/torchcde) packages, which included the CubicSpline used to construct the continuous input sequence (X), and the adaptive step size solver used to solve Eq. [4]. As $l^1_\varnothing$ merely initializes the hidden state, it is a single-layer perceptron. $f_\varnothing$ is a multi-layer perceptron consisting of 6 layers with variable widths. $l^2_\varnothing$ contains multiple parallel multi-layer perceptrons, one for each QMRI parameter (Kaandorp et al., 2021). For further details on the architecture of the NCDEs, see Table S1.

We trained an NCDE to map the input sequence to M QMRI parameters, where $M = 2$ for VFA *T1*-mapping, $M = 4$ for IVIM MRI and $M = 4$ for DCE-MRI. Results on parameters required for modelling but not representing relevant physiological quantities (s_0 in VFA *T1*-mapping and IVIM MRI, $d\tau$ in DCE-MRI) are not presented. To constrain the estimates, QMRI parameters were rescaled to pre-specified ranges [$y_{m,min}$, $y_{m,max}$] using:

$$\hat{y}_m = y_{m,min} + \sigma\left(\hat{y}_{m,coeff}\right)\left(y_{m,max} - y_{m,min}\right), \tag{5}$$

where $\hat{y}_m$ is the estimate for y_m (m is an integer in range: [1,M]), $\hat{y}_{m,coeff}$ is the NCDE output ($\hat{y}_{m,coeff}$ has range $[-\infty, \infty]$) and σ is the sigmoid

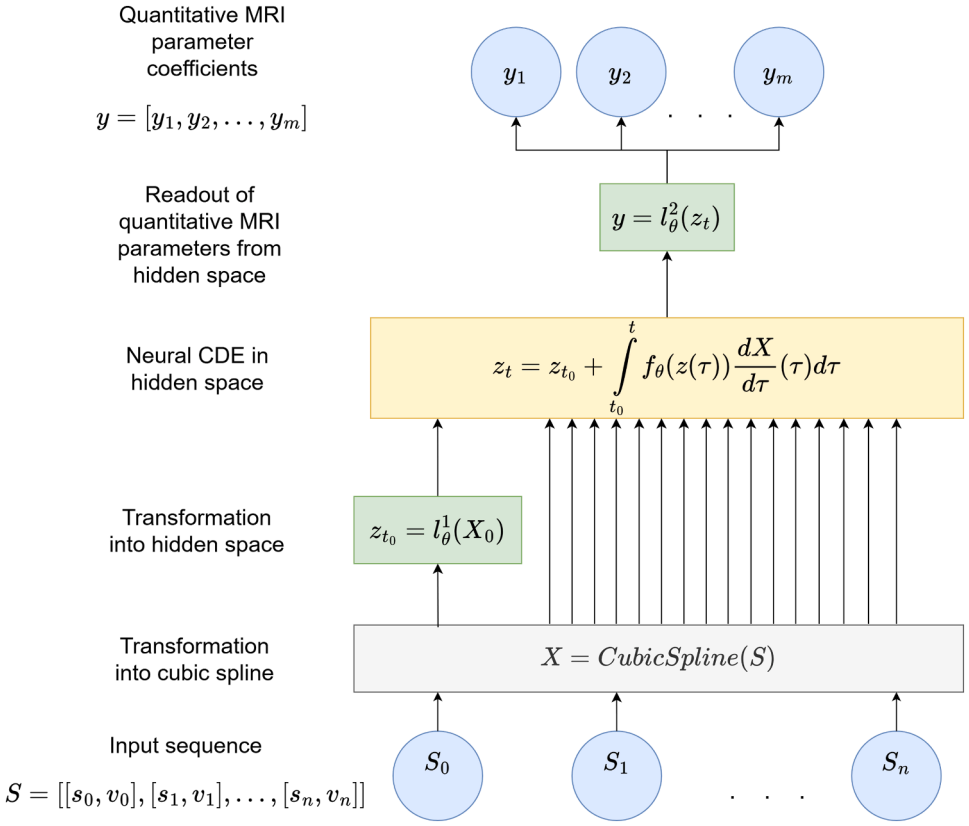

Fig. 1. Schematic representation of an NCDE for QMRI parameter estimation. The NCDE is composed of three neural networks: $\boldsymbol{l^1_\varnothing}$, $\boldsymbol{f_\varnothing}$, and $\boldsymbol{l^2_\varnothing}$. The input sequence $\boldsymbol{S}$ can have arbitrary length and irregular sampling intervals. The length of output vector $\boldsymbol{y}$ corresponds to the number of predicted QMRI parameter coefficients. $\boldsymbol{X}$ is a twice continuously differentiable cubic spline, whose knots are the elements of $\boldsymbol{S}$. $\boldsymbol{z}$ denotes the hidden state.

function, mapping the parameter-specific output of the NCDE to normalized range $[0,1]$. The ranges for rescaling factors $y_{m,min}$ and $y_{m,max}$ are described per QMRI parameter in Table 1. Based on the estimated QMRI parameters together with the set of measured values for the independent variable (flip angle in VFA *T1*-mapping, b-value in IVIM MRI and time in DCE-MRI), a signal curve S_{pred} is estimated according to Eqs. (1-3).

For each QMRI forward model, denoted in Eqs. (1-3), a dedicated NCDE is trained by minimizing a loss (L), which is a linear combination of a supervised loss and a physics-informed loss (Barbieri et al., 2020), following:

$$L = \frac{1}{len\left(\hat{y}_{m,coeff}\right)} \sum_{m=1}^{len\left(\hat{y}_{m,coeff}\right)} \left(\sigma\left(\hat{y}_{m,coeff}\right) - \sigma\left(y_{m,coeff}\right)\right)^2 + \frac{1}{len\left(s_{n,input}\right)} \sum_{n=1}^{len\left(s_{n,input}\right)} \left(s_{n,input} - s_{n,pred}\right)^2, \quad (6)$$

where $\sigma\left(\hat{y}_{m,coeff}\right)$ and $\sigma\left(y_{m,coeff}\right)$ represent the m-th elements of range-normalized QMRI parameter estimates and range-normalized ground truth QMRI parameters, respectively. $s_{n,input}$ and $s_{n,pred}$ represent the n-th elements of the normalized input signal and the estimated signal curves, respectively. Note that n typically changes per input signal. When trained on in vivo data, where no supervisory coefficients for the estimated QMRI parameters are available, the combined loss in Eq. (6) reduces to the physics-informed loss.

2.2.2. Multi-layer perceptron model

As a reference, we trained conventional acquisition-specific multi-layer perceptron (MLP) models to estimate QMRI parameters. Acquisition-specific MLPs learn a parameterized mapping from the input sequence S to output y. The input sequence (S) contained the signal intensities (s, either a series of MRI signals for VFA *T1*-mapping and IVIM MRI or a derived gadolinium concentration curve for DCE-MRI).

In our implementation, we kept the architecture similar to the described implementation of IVIM-NET (Kaandorp et al., 2021), which contains multiple parallel MLPs, one for each QMRI parameter. This corresponds to the described implementation of $l^2_{\emptyset}$ in the NCDE model (Fig. 1). For further details on the architecture of the MLPs, see Table S2.

Equal to the method we used to estimate parameters using NCDEs, we trained each MLP to map the input sequence to M QMRI parameters, where $M = 2$ for VFA *T1*-mapping, $M = 4$ for IVIM MRI and $M = 4$ for DCE-MRI. Results on parameters required for modelling but not representing relevant physiological quantities (s_0 in VFA *T1*-mapping and IVIM MRI, $d\tau$ in DCE-MRI) are not presented. To constrain the estimates, QMRI parameters were rescaled to pre-specified ranges using Eq. (5).

For each QMRI forward model, denoted in Eqs. (1-3), a dedicated MLP is trained by minimizing a loss, which is a linear combination of a physics-informed loss (Barbieri et al., 2020) and a supervised loss, described in Eq. (6).

Table 1
Description of QMRI parameter ranges used in simulated data during training and evaluation. QMRI parameters were randomly sampled from a uniform distribution with these ranges.

Quantitative MRI model	Parameter	Range
VFA *T1*-mapping	*T1*	100 – 3000 ms
IVIM MRI	D	0.00035 – 0.003 mm^2/s
	D^*	0.05 – 0.1 mm^2/s
	f	0.03 – 0.25
DCE-MRI	k_e	0.1 – 2 min^{-1}
	v_e	0.01 – 0.7
	v_p	0.001 – 0.05

When trained on in vivo data, where no supervisory coefficients for the estimated QMRI parameters are available, the combined loss in Eq. (6) reduces to the physics-informed loss.

Note that as each MLP is trained for 1 acquisition setting, over the following experiments we trained many MLPs, one for each particular acquisition setting.

2.3. Data

2.3.1. Simulations

To show the generalizability of NCDEs in the context of QMRI parameter estimation, we trained three different NCDE models and three different acquisition-specific MLPs to estimate IVIM MRI, VFA *T1*-mapping and DCE-MRI parameters.

Per QMRI model, 1,000,000 training signal (for IVIM MRI and VFA *T1*-mapping) or concentration (for DCE-MRI) curves were simulated using the forward model as described in Eqs. (1-3). Individual QMRI parameters were sampled from a uniform distribution with ranges as described in Table 1.

Signals generated from acquisition protocols with varying length and configurations of the independent variable were used for training and evaluation of the NCDEs and for the evaluation of LSQ fitting. The independent variables (flip angle, b-value or time) were sampled as described in Table 2. The sampling scheme reflects clinical imaging protocols, while maintaining a source of variability in both sequence length and values of the independent variable. This allowed training and evaluating NCDEs and evaluating LSQ fitting on a large range of representative configurations of independent variables.

Signals generated from acquisition protocols with fixed length and configuration of the independent variable were used for training and evaluation of the MLPs as well as for evaluation of the trained NCDEs and LSQ fitting. The independent variables (flip angle, b-value or time) were sampled as described in Table 3.

For VFA *T1*-mapping and IVIM MRI, noise was added to the simulated signal curves to make the data follow a Rician distribution similar to MRI magnitude images. The signal-to-noise ratio was set to a predefined value ($SNR_{predefined}$) and the noisy signal (s_{noisy}) followed:

$$s_{noisy} = \sqrt{\left(s + N\left(0, \frac{s_0}{SNR_{predefined}}\right)\right)^2 + N\left(0, \frac{s_0}{SNR_{predefined}}\right)^2}, \quad (7)$$

where s_0 refers to the signal at $b = 0$ s/mm^2 for IVIM MRI and the magnetization at thermal equilibrium for VFA *T1*-mapping. Here, $N(\mu, SD)$ represents random sampling from a normal distribution with mean μ and standard deviation SD. For DCE-MRI, the measurements reflect gadolinium concentration instead of an MR signal and therefore Gaussian noise instead of Rician noise was added to the generated curves following:

$$s_{noisy} = s + N\left(0, \frac{s_{max}}{SNR_{predefined}}\right). \quad (8)$$

In DCE-MRI, s_{max} refers to the max signal of the signal curve.

2.3.2. In vivo

2.3.2.1. VFA T1-mapping. VFA *T1*-mapping MRI data of the abdomen, leg and brain were acquired in 8 healthy volunteers using a 3T MRI scanner (Ingenia, Philips, Best, The Netherlands) at the Amsterdam UMC. By scanning three anatomical regions, a broad range of *T1* values was covered in the experiments. In one volunteer, no data from the abdomen was acquired. In another volunteer, no data from the leg was acquired. The study was approved by the local medical ethics review committee. All participants provided written informed consent, and research was performed in accordance with the Declaration of Helsinki guidelines. To obtain a complete VFA *T1*-mapping dataset, *T1*-weighted

Table 2
Description of the protocol for sampling the independent variables and simulating the corresponding signal curves. These simulation procedures have been used during training and evaluation. Even though the denoted range in SNR appears vastly different among the three QMRI simulations, this is caused by differences in the SNR definition. The ceiling operator ($\lceil \boldsymbol{x} \rceil$) rounds $\boldsymbol{x}$ up to the nearest integer, the floor operator ($\lfloor \boldsymbol{x} \rfloor$) rounds $\boldsymbol{x}$ down to the nearest integer.

Quantitative MRI model	Sequence length	Indepen-dent variable	Range of independent variable	Sampling protocol	SNR range
VFA *T1*-mapping	[3, 8]	Flip angle	[0, 30]°	For N flip angles: 1. Sample flip angle of 1° 2. Randomly sample 1 flip angle in range $[3, 10]$° 3. Randomly sample 1 flip angle in range $[15, 30]$° 4. Equidistantly sample $N-3$ flip angles between 1° and maximum flip angle	[50, 300]
IVIM MRI	[4, 10]	b-value	[0, 800] s/mm^2	For N b-values: 1. Sample b-value of 0 s/mm^2 2. Randomly sample $\left\lfloor \frac{\left\lceil \left(\frac{N-1}{2}\right) \right\rceil}{2} \right\rfloor$ b-values <50 s/mm^2 3. Randomly sample $\left\lfloor \frac{\left\lceil \left(\frac{N-1}{2}\right) \right\rceil}{2} \right\rfloor$ b-values in range $[50, 99]$ s/mm^2 4. Randomly sample $\left\lfloor \left(\frac{N-1}{2}\right) \right\rfloor$ b-values [100, 800] s/mm^2	[5, 40]
DCE-MRI	[15, 20]	Time	[0, 240] s	For N timepoints: 1. Randomly sample temporal resolution in range $[8, 12]$ s	[5, 40]

3D spoiled gradient echo MRI was acquired at 30 flip angles (1-30°). Detailed MRI acquisition settings are provided in Table S2.

2.3.2.2. IVIM MRI. IVIM MRI data of the abdomen was acquired in 11 patients using a 3T MRI scanner (Ingenia, Philips, Best, The Netherlands) at the Amsterdam UMC. The study was approved by the local medical ethics review committee. All participants provided written informed consent, and research was performed in accordance with the Declaration of Helsinki guidelines. Detailed MRI acquisition settings are provided in already published work (Gurney-Champion et al., 2018).

2.3.3. Preprocessing

Neural networks perform optimally for normalized signals, while in VFA *T1*-mapping and IVIM MRI signals are in arbitrary units. Therefore, we normalized the data by dividing the signal curve by an approximate s_0.

In VFA *T1*-mapping, at low flip angles Eq. (1) can be rewritten to:

$$s_0 \approx \frac{s(FA)}{\sin(FA)}. \tag{9}$$

To get a reliable estimate of s_0, a measurement at 1° was always included in the sampled sequences during training and evaluation. In IVIM MRI, s_0 was approximated by the signal at b=0 s/mm^2. For DCE-MRI modelling, the input signal is the contrast concentration in M (molarity), and is thus inherently normalized.

Table 3
Description of the sampled independent variables for training of the MLPs and for the fixed-acquisition evaluations of the MLPs, NCDEs and LSQ fitting.

Data type	QMRI application	Sequence length used in training and evaluation	Sampled independent variables
Simulation	VFA *T1*-mapping	4	Flip angles: [1, 6, 10, 25]°
	IVIM MRI	7	B-values: [0, 10, 20, 50, 100, 300, 600] s/mm^2
	DCE-MRI	17	Timepoints: [9, 18, 27, 36, 45, 54, 63, 72, 81, 90, 99, 108, 117, 126, 135, 144, 153] s
In vivo	VFA *T1*-mapping	3	Flip angles: [1, 6, 29]°
		4	Flip angles: [1, 6, 15, 29]°
		5	Flip angles: [1, 6, 10, 20, 29]°
	IVIM MRI	4	B-values: [0, 10, 50, 600] s/mm^2
		6	B-values: [0, 10, 20, 50, 75, 600] s/mm^2
		8	B-values: [0, 10, 20, 50, 75, 150, 400, 600] s/mm^2
		10	B-values: [0, 10,20, 30, 50, 75, 150, 250, 400, 600] s/mm^2

2.4. Experiments

2.4.1. Training and evaluation on simulation data

2.4.1.1. NCDE training and evaluation. Three NCDE models (one for each QMRI forward model) were trained for 750 epochs, with each epoch consisting of 100 data batches of size 64. Empirically, we observed that training for 750 epochs with batch size of 64 and 100 steps per epoch led to converged models across the different estimation methods and QMRI forward models. In order to keep the methodology among different estimation methods as similar as possible, we opted to train all models for the same amount of iterations. Each training set (separate for VFA *T1*-mapping, IVIM MRI and DCE-MRI) contained 1,000,000 simulated signal curves. A validation set was not used, as the training was based on a fixed number of epochs. These curves were generated with varying sequence lengths and configurations of independent variables, following the procedure outlined in Section 2.3.1 and Table 2. Each training used the Adam optimizer (Kingma et al, 2015). Learning rate started at $1{\cdot}10^{-4}$, and halved at epochs 250, 350, 450, 550 and 650. During training, an adaptive step solver adjusts the step size so that the error in the solution remains approximately equal to a predefined tolerance (Kidger, 2020). Absolute and relative error tolerance of the adaptive ODE solver to solve for the hidden state in Eq. (4) were set to 0.00001 and 0.001 respectively and halved at epochs 300, 400, 500 and 600.

For the comparison between LSQ fitting and NCDEs, QMRI

parameter estimates were evaluated across different sequence lengths and SNR levels. To achieve this, 1,000,000 curves per QMRI forward model (VFA *T1*-mapping, IVIM MRI and DCE-MRI) were generated with varying sequence lengths and configurations of independent variables, following the procedure outlined in Section 2.3.1 and Table 2.

For the comparison among MLPs, NCDEs and LSQ fitting, the NCDEs trained on varying sequence lengths and configurations of the independent variable were evaluated on simulated signal curves sampled according to the acquisition protocol described in Table 3.

Since for simulations the ground truth QMRI parameter values are known, the error (as in Eq. (10)) was calculated using:

$$\text{error} = \hat{y} - y. \tag{10}$$

Where $\hat{y}$ denotes the estimated QMRI parameter and y the ground truth QMRI parameter. The squared error was calculated using:

$$\text{squared error} = (\hat{y} - y)^2. \tag{11}$$

In addition, the mean, median, 25th and 75th percentiles of the squared error over all curves were calculated. The mean of the squared error functions as a measure of accuracy. The mean of the error and the variance of the error over all curves was calculated. The mean of the error functions as a measure of bias, the variance of the error functions as a measure of precision.

2.4.1.2. MLP training and evaluation. Three MLPs (one for each forward model) were trained for 750 epochs, with each epoch consisting of 100 data batches of size 64. Each training set (separate for VFA *T1*-mapping, IVIM MRI and DCE-MRI) contained 1,000,000 simulated signal curves. A validation set was not used, as the training was based on a fixed number of epochs. The simulated signal curves used for the training of the MLPs followed the acquisition protocol described in Table 3. Each training used the Adam optimizer (Kingma et al., 2015). Learning rate started at $1{\cdot}10^{-4}$, and halved at epochs 250, 350, 450, 550 and 650.

QMRI parameter estimates were evaluated across SNR levels at a predefined acquisition protocol and subsequently compared to LSQ fitting and NCDE-based parameter estimates. To achieve this, 1,000,000 curves per QMRI forward model (VFA *T1*-mapping, IVIM MRI and DCE-MRI) were generated following the acquisition protocol described in Table 3. Since for simulations the ground truth QMRI parameter values are known, the error was calculated as in Eq. 10 and the squared error was calculated as in Eq. 11.

In addition, the mean, median, 25th and 75th percentiles of the error and squared error over all curves were calculated. The mean of the squared error functions as a measure of accuracy. The mean of the error and the variance of the error over all curves was calculated. The mean of the error functions as a measure of bias, the variance of the error functions as a measure of precision.

2.4.2. Training and evaluation on in vivo data

2.4.2.1. NCDE training and evaluation. Eight NCDEs (one for each of four cross-validation folds times one for each in vivo QMRI forward model) were trained to estimate QMRI parameters using 4-fold cross-validation, with each fold utilizing data from 75% of the available datasets for training, while leaving out 25% of data for evaluation. All models used the same training hyperparameters. Each training set (specific to each fold) was supplemented with simulated signal curves, to ensure that all possible QMRI parameter combinations were seen during training. For NCDE training, both the in vivo and simulated signal curves were generated with varying sequence lengths and configurations of independent variables, following the procedure outlined in Section 2.3.1 and Table 2. In each batch, in vivo signal curves were sampled with a 6:1 probability compared to simulated signal curves. Training lasted for 750 epochs, with each epoch consisting of 100 data batches of size 64. The model trained on simulation data of the respective QMRI forward model was taken as initialization. The learning rate started at $2.5{\cdot}10^{-5}$, and halved at epochs 250, 350, 450, 550 and 650. Optimization was performed using the Adam optimizer (Kingma et al., 2015). Absolute and relative error tolerance of the adaptive step size solver, used to compute the hidden state in Eq. (4) were set to 0.00001 and 0.001 respectively and halved at epochs 300, 400, 500 and 600.

For the comparison to LSQ fitting, both the in vivo and simulated signal curves were generated with varying sequence lengths and configurations of independent variables, following the procedure outlined in Section 2.3.1 and Table 2. For the comparison to MLPs, the NCDEs trained on varying sequence lengths and configurations of the independent variable were evaluated on in vivo signal curves sampled according to the acquisition protocol described in Table 3.

2.4.2.2. MLP training and evaluation. Twenty-eight acquisition-specific MLPs (one for each of four cross-validation folds times three acquisition schemes for VFA *T1*-mapping and four acquisition schemes for IVIM MRI) were trained to estimate QMRI parameters using 4-fold cross-validation, with each fold utilizing data from 75% of the available datasets for training, while leaving out 25% of data for evaluation. All models used the same training hyperparameters. Each training set (specific to each fold) was supplemented with simulated signal curves, to ensure that all possible QMRI parameter combinations were seen during training. For MLP training, the in vivo and simulated signal curves followed the acquisition protocol described in Table 3. In each batch, in vivo signal curves were sampled with a 6:1 probability compared to simulated signal curves. Training lasted for 750 epochs, with each epoch consisting of 100 data batches of size 64. The model trained on simulation data of the respective QMRI forward model was taken as initialization. The learning rate started at $2.5{\cdot}10^{-5}$, and halved at epochs 250, 350, 450, 550 and 650. Optimization was performed using the Adam optimizer (Kingma et al., 2015).

For the comparison to NCDEs and LSQ fitting, the MLPs were evaluated on in vivo signal curves sampled according to the acquisition protocol described in Table 3.

2.4.2.3. VFA T1-mapping. Assessing the accuracy of QMRI parameter estimation methods in vivo is challenging due to the absence of ground truth for QMRI parameters. To establish reference values, an LSQ fit using all 30 flip angles was used as the reference standard. LSQ fitting and NCDE-based parameter estimations were performed on input sequences containing varying subsets of flip angles (3, 4 and 5 flip angles), which were generated according to the sampling protocol in Table 2. Within each patient, the sampling of different sequence lengths was initialized with the same random seed to ensure similarity across subsets of different length. Also, MLP-based parameter estimation, LSQ fitting and NCDE-based parameter estimations were performed on input sequences containing fixed subsets of flip angles (3, 4 and 5 flip angles), which were generated according to the sampling protocol in Table 3. Errors (as defined in Eq. (10)) and squared errors (as defined in Eq. (11)) between the reference standard and estimates were calculated per signal curve within a foreground mask, which excluded data with only zero values. The mean error and mean squared error were computed for each volunteer and anatomical region (abdomen, brain, and leg), and box plots were used to visualize the distribution of these metrics, grouped by anatomical region.

2.4.2.4. IVIM MRI. In contrast to the application to VFA *T1*-mapping, establishing general reference values applicable to all estimation methods (NCDE, MLP and LSQ fitting respectively) was not feasible. Instead, we used method-specific (NCDE, MLP and LSQ fitting respectively) reference values with 10 b-values, and assessed the consistency of parameter estimates over various subsets of the input sequence. LSQ fitting and NCDE-based parameter estimations were performed on input

sequences containing varying subsets of b-values (4, 6 and 8 b-values), which were generated according to the sampling protocol in Table 2. Within each patient, the sampling of different sequence lengths was initialized with the same random seed to ensure similarity across subsets of different length. Also, MLP-based parameter estimation, LSQ fitting and NCDE-based parameter estimations were performed on input sequences containing fixed subsets of b-values (4, 6 and 8 b-values), which were generated according to the sampling protocol in Table 3. Errors (as defined in Eq. (10)) and squared errors (as defined in Eq. (11)) between the reference standard and estimates were calculated per signal curve within a foreground mask, which excluded data with no apparent signal decay over b-values. The mean error and mean squared error were computed for each patient, and box plots were used to visualize the distribution of these metrics.

3. Results

3.1. Simulations

3.1.1. VFA T1-mapping

Across all SNR levels < 200, NCDE-based parameter estimation consistently had a lower mean squared error, lower median (except for SNR $= 150$) and lower 75th percentile of the squared error than LSQ fitting at estimating the relaxation constant $T1$ (Fig. 2, left). However, at the highest SNR levels (SNR ≥ 200), LSQ fitting performed better than NCDE-based parameter estimation in estimating $T1$. Both NCDEs and LSQ fitting had more accurate estimates when the input sequence became longer (Fig. 2, center). For all sequence lengths, NCDEs had a lower mean squared error than LSQ fitting in estimating the relaxation constant $T1$. When evaluated at a fixed acquisition scheme, both NCDEs and MLPs had a lower mean squared error than LSQ fitting at all SNR $\leq$ 200 (Fig. 2, right). While MLPs have lower mean squared errors than NCDEs at low SNR levels, this effect disappears at higher SNR. NCDE-based and MLP-based parameter estimation yielded a decrease in variance of the $T1$ error compared to LSQ fitting across all SNR levels, as illustrated in Fig S1. This reduction in variance of the $T1$ error led to an increase in bias of the $T1$ error, especially at low SNR levels. NCDEs estimated QMRI parameters at a rate of 310 VFA $T1$-mapping curves per second, compared to 277 curves per second for LSQ fitting and 714 curves per second for MLPs. Training time was 63 seconds per epoch for NCDEs, compared to 4 seconds per epoch for MLPs.

3.1.2. IVIM MRI

At low SNR levels (SNR ≤ 20), NCDEs had lower mean squared error and median of the squared error than LSQ fitting across all IVIM MRI parameters (Fig. 3, middle row). At high SNR levels (SNR ≥ 40), LSQ improved estimates of the diffusion coefficient (D) and perfusion fraction (f) compared to NCDEs, although NCDEs retained better estimates of the pseudo-diffusion coefficient (D^*). Both NCDEs and LSQ fitting had more accurate estimates when the input sequence became longer (Fig. 3, bottom row). For almost all sequence lengths, NCDEs had lower mean squared error and lower 75th percentile of the squared error than LSQ fitting in estimating IVIM MRI parameters. When evaluated at a fixed acquisition scheme, both NCDEs and MLPs had a lower mean squared error than LSQ fitting at all SNR ≤ 30 for all parameters, except for the diffusion coefficient (D) at SNR 10 where the MLP has a higher mean squared error than LSQ and the perfusion fraction (f) at SNR 20 where the NCDE has a higher mean squared error than LSQ (Fig. 3, top row). NCDE-based and MLP-based parameter estimation yielded a decrease in variance of the errors in IVIM parameters compared to LSQ fitting across all SNR levels, as illustrated in Fig S2. In the case of NCDEs, this reduction in variance of the error led to an increase in bias of the error at high SNR levels for all IVIM parameters. NCDEs estimated QMRI parameters for 133 IVIM curves per second, compared to 140 curves per second for LSQ fitting and 500 curves per second for MLPs. Training time was 107 seconds per epoch for NCDEs, compared to 5 seconds per epoch for MLPs.

3.1.3. DCE-MRI

At almost all SNR levels (SNR ≥ 10), NCDEs had a lower mean squared error than LSQ fitting across all DCE-MRI parameters (Fig. 4, middle row). In almost all of those cases, NCDEs also showed a lower 75th percentile of the squared error. At very low SNR (SNR $= 5$), LSQ had a lower mean squared error than NCDEs in estimating the reflux rate ($\boldsymbol{k_e}$) and the fractional volume of the extravascular extracellular space ($\boldsymbol{v_e}$), while NCDEs had a lower mean squared error in estimating the fractional volume of plasma ($\boldsymbol{v_p}$). Both NCDEs and LSQ fitting had more accurate estimates when the input sequence became longer (Fig. 4, bottom row). For almost all sequence lengths, NCDEs had a lower mean squared error and lower 75th percentile of the squared error than LSQ fitting in estimating the DCE-MRI parameters. When evaluated at a fixed acquisition scheme, both NCDEs and MLPs had a lower mean squared error than LSQ fitting at all SNR ≤ 40 for all parameters, except for the lowest SNR of 5 where the NCDE has a higher mean squared error than LSQ and MLP for estimating the reflux rate ($\boldsymbol{k_e}$) and the fractional

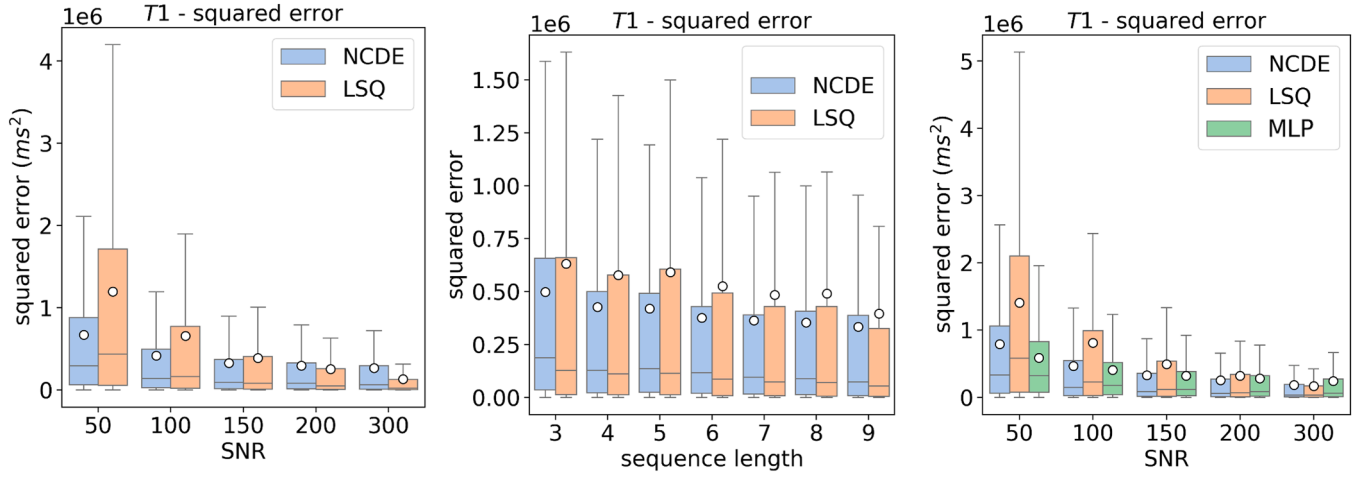

Fig. 2. Left: box plots of squared errors of estimated $T1$ parameters as a function of SNR for NCDE (blue) and LSQ (orange). Center: box plots of squared errors of estimated $T1$ parameters as a function of sequence length for NCDE (blue) and LSQ (orange). Acquisition-specific MLPs cannot be evaluated at variable acquisition schemes (left and center image). Right: box plots of squared errors of estimated $T1$ parameters as a function of SNR for NCDE (blue), LSQ (orange) and MLP (green), evaluated at a fixed acquisition scheme ([1, 6, 10, 25]°). The white circle represents the mean squared error.

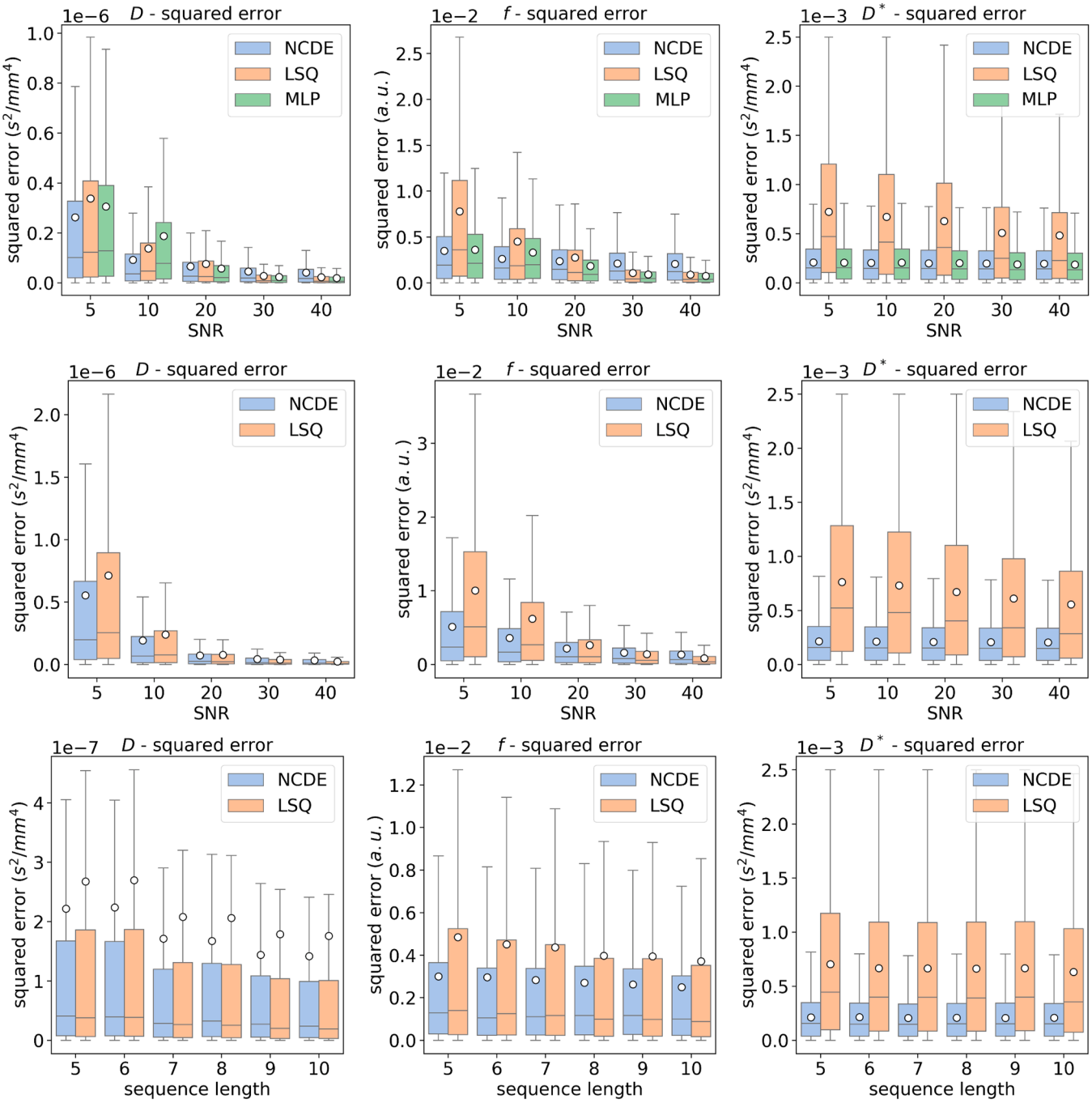

Fig. 3. Top row: box plots of squared errors of estimated IVIM parameters as a function of SNR for NCDE (blue), LSQ (orange) and MLP (green), evaluated at a fixed acquisition scheme ([0, 10, 20, 50, 100, 300, 600] s/mm^2). Middle row: box plots of squared errors of estimated IVIM parameters as a function of SNR for NCDE (blue) and LSQ (orange). Bottom row: box plots of squared errors of estimated IVIM parameters as a function of sequence length for NCDE (blue) and LSQ (orange). Acquisition-specific MLPs cannot be evaluated at variable acquisition schemes (middle and bottom rows). The white circle represents the mean squared error.

volume of the extravascular extracellular space ($\boldsymbol{v_e}$) (Fig. 4, top row). NCDE-based and MLP-based parameter estimation yielded a decrease in variance of the errors in DCE parameters compared to LSQ fitting across all SNR levels, except for SNR $\leq$ 10 in the reflux rate ($\boldsymbol{k_e}$) error and for SNR = 5 in the extravascular extracellular space ($\boldsymbol{v_e}$) error, as illustrated in Fig S3. In the case of NCDEs, this reduction in variance led to an increase in bias of the extravascular extracellular space ($\boldsymbol{v_e}$) error at high SNR levels, and an increase in bias of the reflux rate ($\boldsymbol{k_e}$) error at low SNR levels. NCDEs estimated QMRI parameters for 60 DCE-MRI curves per second, compared to 27 curves per second for LSQ fitting and 303 curves per second for MLPs. Training time was 198 seconds per epoch for NCDEs, compared to 14 seconds per epoch for MLPs.

3.1.4. *VFA T1-mapping*

NCDE-based parameter estimation resulted in more accurate QMRI parameter maps in vivo, especially in the abdomen and leg. Fig. 5 shows typical examples of the QMRI parameter maps from NCDE-based parameter estimation, MLP-based parameter estimation and LSQ fitting of the abdominal data. Figs. S4 & S5 show additional examples from brain and leg data. In the abdomen, the anatomical region where *T1* values are hardest to estimate, the NCDE better approximated the reference *T1* values in noisy parts of the image than LSQ (Fig. 5, at

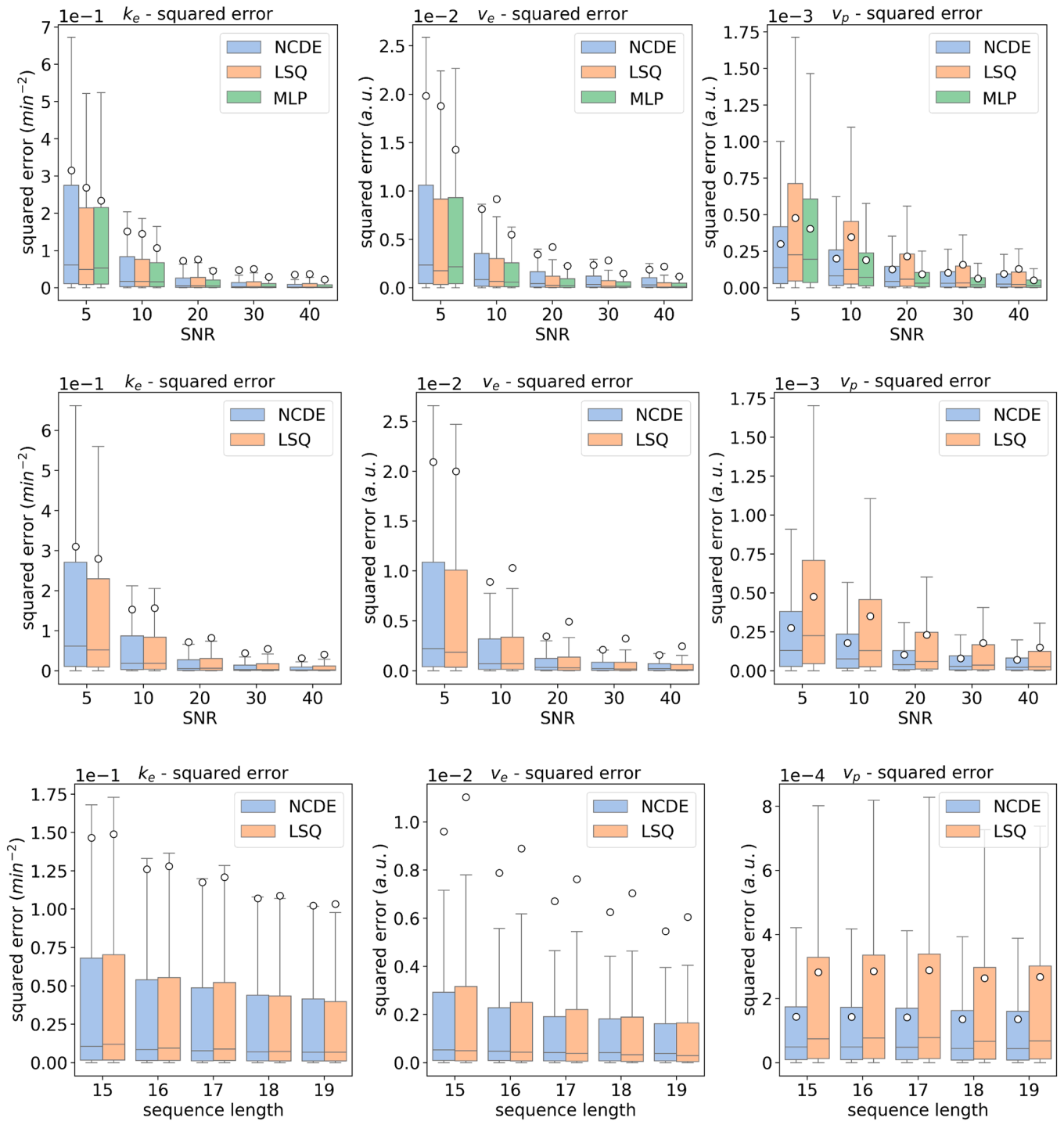

Fig. 4. Top row: box plots of squared errors of estimated DCE parameters as a function of SNR for NCDE (blue), LSQ (orange) and MLP (green), evaluated at a fixed acquisition scheme ([9, 18, 27, 36, 45, 54, 63, 72, 81, 90, 99, 108, 117, 126, 135, 144, 153] *s*). Middle row: box plots of squared errors of estimated DCE parameters as a function of SNR for NCDE (blue) and LSQ (orange). Bottom row: box plots of squared errors of estimated DCE parameters as a function of sequence length for NCDE (blue) and LSQ (orange). Acquisition-specific MLPs cannot be evaluated at variable acquisition schemes (middle and bottom rows). The white circle represents the mean squared error.

arrows). For both NCDE-based parameter estimation and LSQ fitting with a subset of flip angles the delineation of different organs (kidney, liver, renal arteries) became less clear compared to the reference *T1* map. For the MLP-based parameter maps, the contrast decreased between different organs with a reduction in the number of flip angles used. In the anatomical region with the highest SNR, the brain (Fig. S4), the contrast between different tissues (grey matter, white matter, cerebrospinal fluid) was similar among NCDE-based parameter estimation and LSQ fitting with a subset of flip angles and the reference *T1* map. For the legs, the contrast between different tissues (bone, muscle) was similar among NCDE-based parameter estimation and LSQ fitting with a subset of flip angles and the reference *T1* map, but the NCDE-based and MLP-based QMRI parameter maps underestimated the *T1* values in some areas of high *T1* (Fig. S5, at arrows).

In abdominal and leg *T1* maps, NCDE-based parameter estimation had a lower mean, median, 25th and 75th percentile of the mean squared

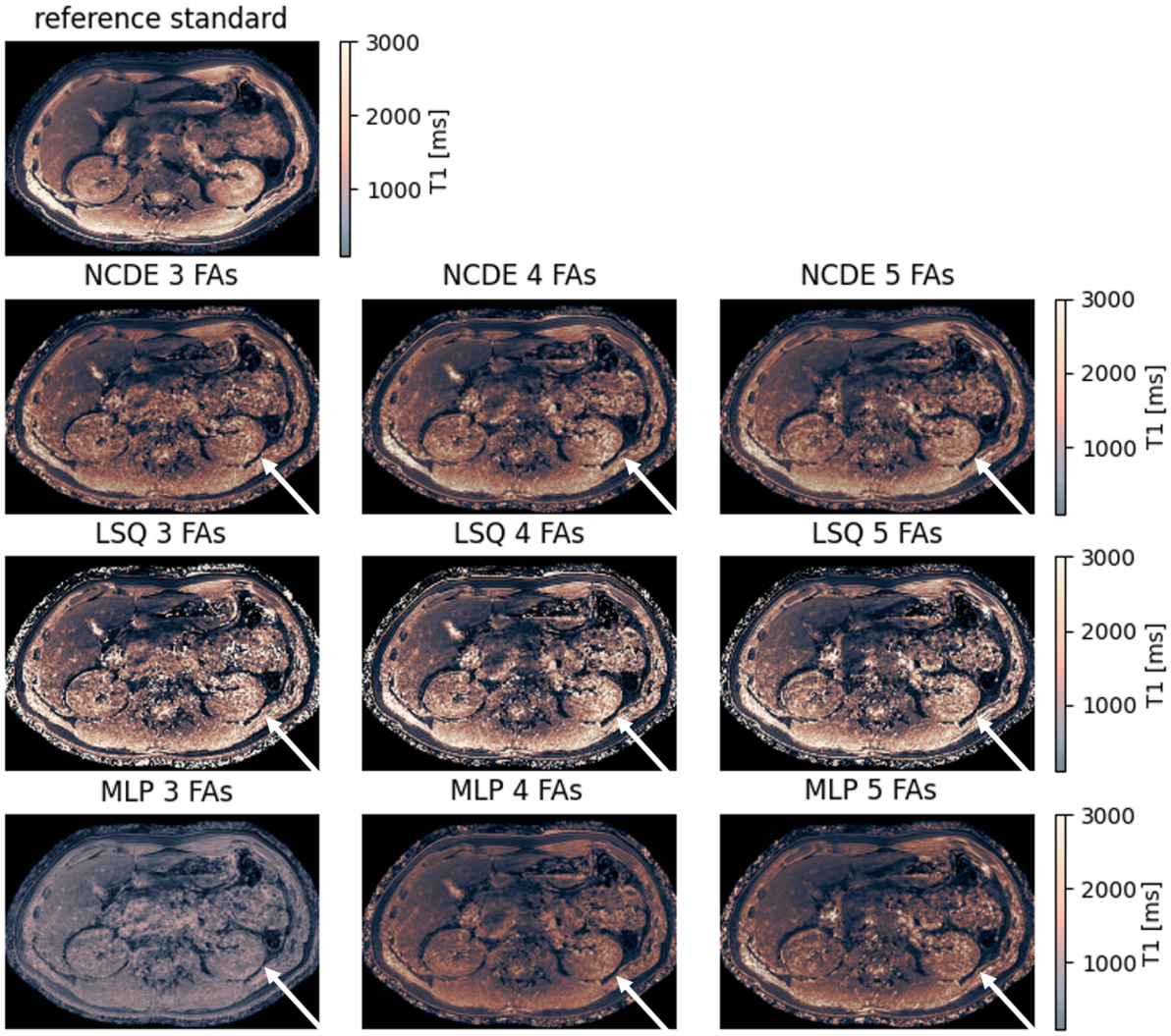

Fig. 5. Comparison of NCDE-based, LSQ-based and MLP-based abdominal *T1* parameter maps based on subsampled data: abdomen reference *T1* parameter map (top row) and estimated *T1* parameter maps for NCDE (2nd row from the top), LSQ (3rd row from the top) and MLP (bottom row). Arrows indicate a region (left kidney) where *T1* parameter maps were degraded by noise, affecting LSQ fitting parameter maps more than NCDE-based parameter maps. MLP-based *T1* parameter maps suffer from poor contrast at 3 and 4 flip angles. Flip angles used for the evaluation at 3, 4 and 5 flip angles were [1, 6, 29]°, [1, 6, 15, 29]° and [1, 6, 10, 20, 29]°, respectively.

error per volunteer over 3, 4 and 5 flip angles (Fig. 6, bottom row) than LSQ fitting. In brain *T1* maps, NCDE-based parameter estimation showed lower mean, median and 25th percentile of the mean squared error per volunteer compared to LSQ fitting for 3, 4 and 5 flip angles, while LSQ showed 75th percentile of the mean squared errors per volunteer at 5 flip angles. The NCDE had no problem dealing with varying input sequence length as the squared errors decreased with increasing sequence length. When evaluated at a fixed acquisition scheme, both NCDE-based parameter estimation and MLP-based parameter estimation showed lower mean squared error than LSQ fitting across sequence lengths for abdominal and leg data, while LSQ fitting showed lower mean squared error than NCDE-based parameter estimation and MLP-based parameter estimation for brain data at 4 and 5 flip angles. NCDE-based and MLP-based parameter estimation yielded a smaller bias of the error compared to LSQ fitting across sequence length for data of the abdomen, but a larger bias for brain data, as illustrated in Fig. S8 (top row). NCDE-based parameter estimation yielded a smaller variance of the error compared to LSQ fitting across sequence length and anatomical regions (Fig. S8, bottom row). MLP-based parameter estimation yielded a smaller variance of the error compared to LSQ fitting across sequence length for abdominal and leg data, but a larger variance of the error for brain data (Fig. S8, 2nd row from the top).

3.1.5. IVIM MRI

NCDE-based, MLP-based and LSQ-based parameter estimation resulted in consistent QMRI parameter maps in vivo when evaluated over subsets of b-values of variable length. Fig. 7 shows typical examples of the abdominal perfusion fraction (f) parameter maps from NCDE-based parameter estimation, MLP-based parameter estimation and LSQ fitting. The LSQ-based perfusion fraction (f) parameter maps show larger variation of the perfusion fraction over the different sequence lengths compared to the NCDE-based and MLP-based parameter maps (Fig. 7, at arrows). Fig. S6 & Fig. S7 show typical examples of the abdominal diffusion coefficient (D) and pseudo-diffusion coefficient (D^*) parameter maps from NCDE-based parameter estimation, MLP-based parameter estimation and LSQ fitting.

Compared to the method-specific reference values obtained at 10 b-values, NCDE-based parameter estimation had a lower mean, median, 25th and 75th percentile of the mean squared error per patient over 4, 6 and 8 b-values (Fig. 8) across all IVIM parameters than LSQ fitting, except for the 75th percentile at 8 b-values for the diffusion coefficient

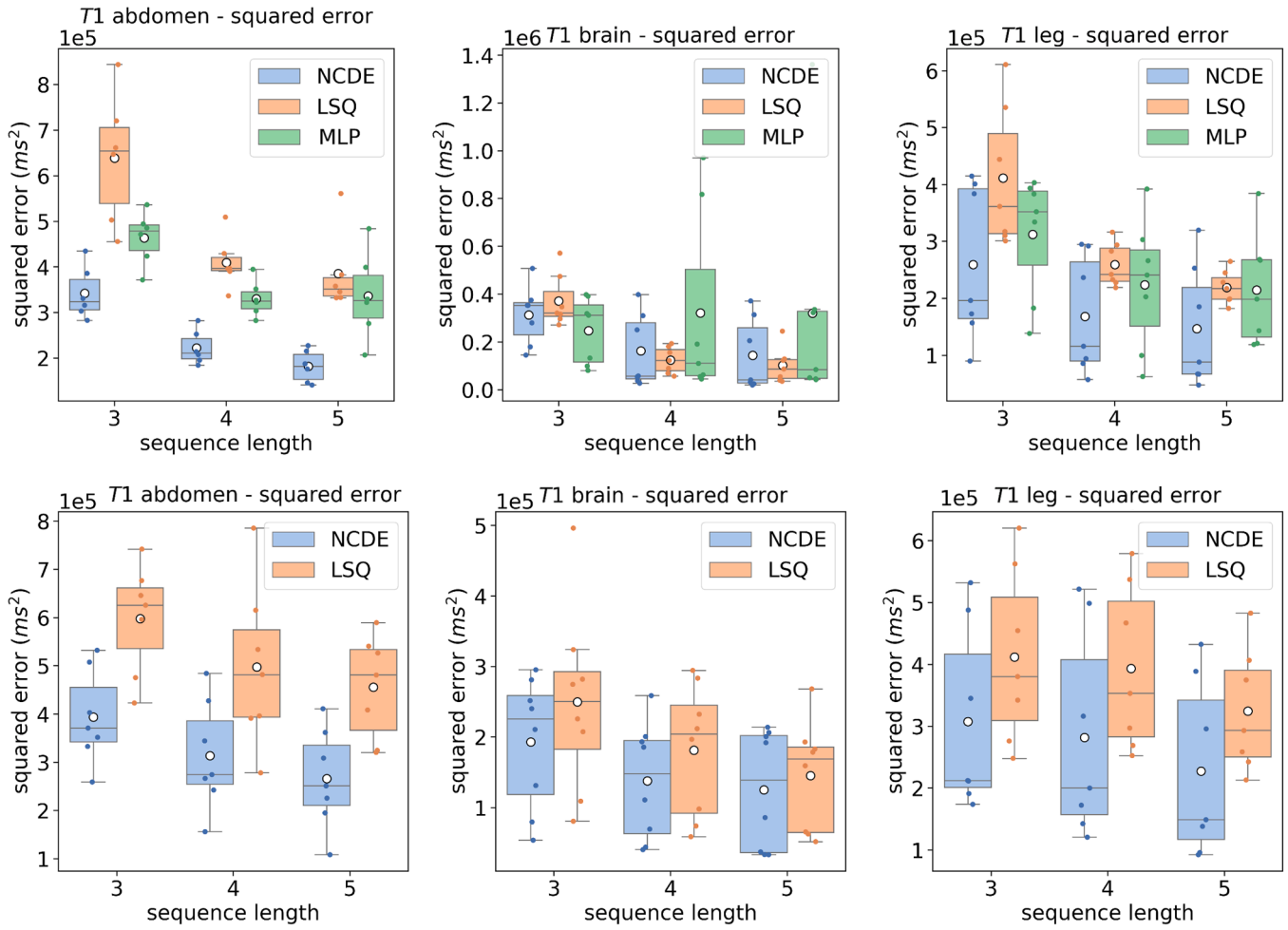

Fig. 6. Top row: box plot of mean squared errors per volunteer of the estimated *T1* parameter by NCDEs (blue), LSQ (orange) and MLP (green) as a function of sequence length, grouped per anatomical region. Flip angles used for the evaluation at 3, 4 and 5 flip angles were [1, 6, 29]°, [1, 6, 15, 29]° and [1, 6, 10, 20, 29]°, respectively. Bottom row: box plot of mean squared errors per volunteer of the estimated *T1* parameter by NCDEs (blue), LSQ (orange) as a function of sequence length, grouped per anatomical region. Acquisition-specific MLPs cannot be evaluated at variable acquisition schemes (bottom row). The white circle represents the mean squared error over all volunteers.

(*D*). When evaluated at a fixed acquisition scheme, both NCDE-based parameter estimation and MLP-based parameter estimation showed lower mean squared error than LSQ fitting across parameters and sequence lengths. NCDE-based parameter estimation showed lower mean, median, 25th and 75th percentile of the mean squared error per patient compared to MLP-based parameter estimation at 4, 6 and 8 b-values for the estimation of the diffusion coefficient (*D*). Conversely, for the estimation of the perfusion fraction (*f*), MLP-based parameter estimation showed lower mean, median, 25th and 75th percentile of the mean squared errors per patient compared to NCDE-based parameter estimation at 4 and 6 b-values, while NCDE-based parameter estimation showed the lowest mean, median, 25th and 75th percentile of the mean squared errors per patient at 8 b-values. The NCDE had no problem dealing with varying amounts of input data as the squared errors decreased with increasing sequence length. NCDE-based and MLP-based parameter estimation yielded a smaller variance of the error compared to LSQ fitting across sequence lengths and IVIM parameters except for MLP-based parameter estimation at 8 b-values, as illustrated in Fig. S9 (2nd row from the top and bottom row). NCDE-based parameter estimation yielded a larger bias of the error compared to LSQ fitting across sequence length and anatomical regions (Fig. S9, 3rd row from the top).

4. Discussion

In this study, we implemented NCDEs as generic acquisition-independent networks for QMRI parameter estimation. Our results indicate that NCDEs can function as a generic tool for the accurate estimation of QMRI parameters, regardless of the QMRI sequence length, the configuration of independent variables, or the QMRI forward model used. To the best of our knowledge, we are the first to propose a general solution for QMRI parameter estimation using deep learning with irregularly sampled data across different QMRI forward models, without retraining for different acquisition protocols. Our NCDE outperformed LSQ fitting for a broad range of settings, especially in the presence of uncertainty (low SNR and shorter input sequences). NCDEs appeared to perform on a similar level as acquisition-specific MLPs, while being able to generalize the demonstrated performance to a broad range of acquisition schemes. These findings suggest that with NCDEs, we have solved one of the main challenges for using deep learning for QMRI parameter estimation in a broader clinical and research setting.

In simulations, NCDEs achieve lower mean squared error than LSQ fitting for all QMRI forward models when SNR levels are low. However, at high SNR, LSQ fitting remains the better method for estimating most QMRI parameters. When NCDEs and MLPs improve parameter estimation compared to LSQ fitting, they tend to do so by reducing the variance in estimation errors. In vivo, NCDEs achieve lower mean squared errors

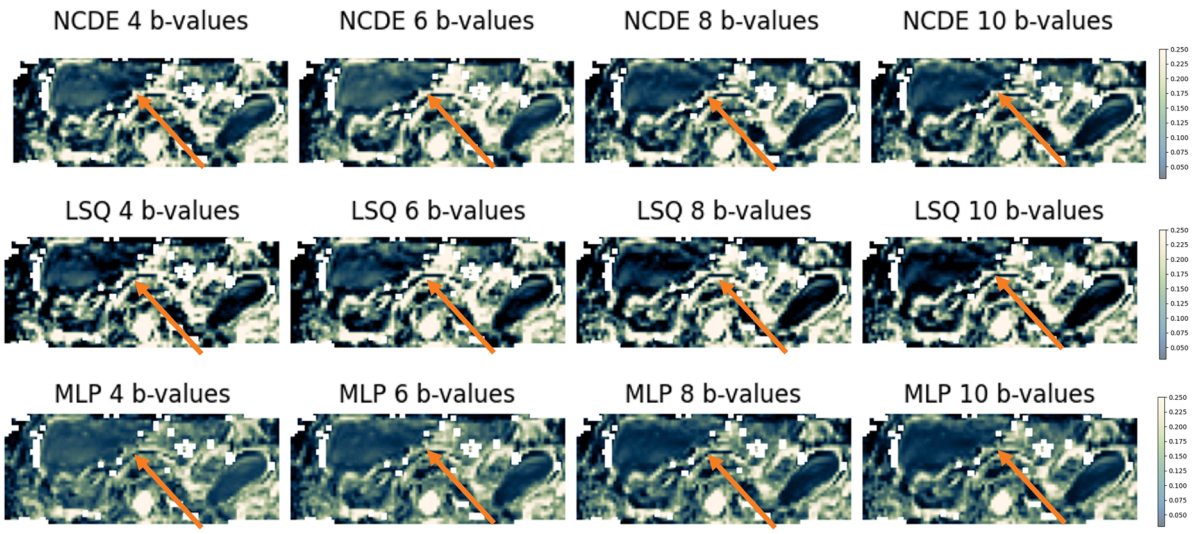

Fig. 7. Comparison of NCDE-based, LSQ-based and MLP-based abdominal perfusion fraction (f) parameter maps based on subsampled data: estimated f parameter maps for NCDE (top row), LSQ (middle row) and MLP (bottom row). Arrows indicate a region (liver) where perfusion fraction parameter maps were inconsistent between different subsets of b-values, affecting LSQ fitting parameter maps more than NCDE-based and MLP-based parameter maps. B-values used for the evaluation at 4, 6, 8 and 10 b-values were [0, 10, 50, 600] s/mm^2, [0, 10, 20, 50, 75, 600] s/mm^2, [0, 10, 20, 50, 75, 150, 400, 600] s/mm^2 and [0, 10, 20, 30, 50, 75, 150, 250, 400, 600] s/mm^2, respectively.

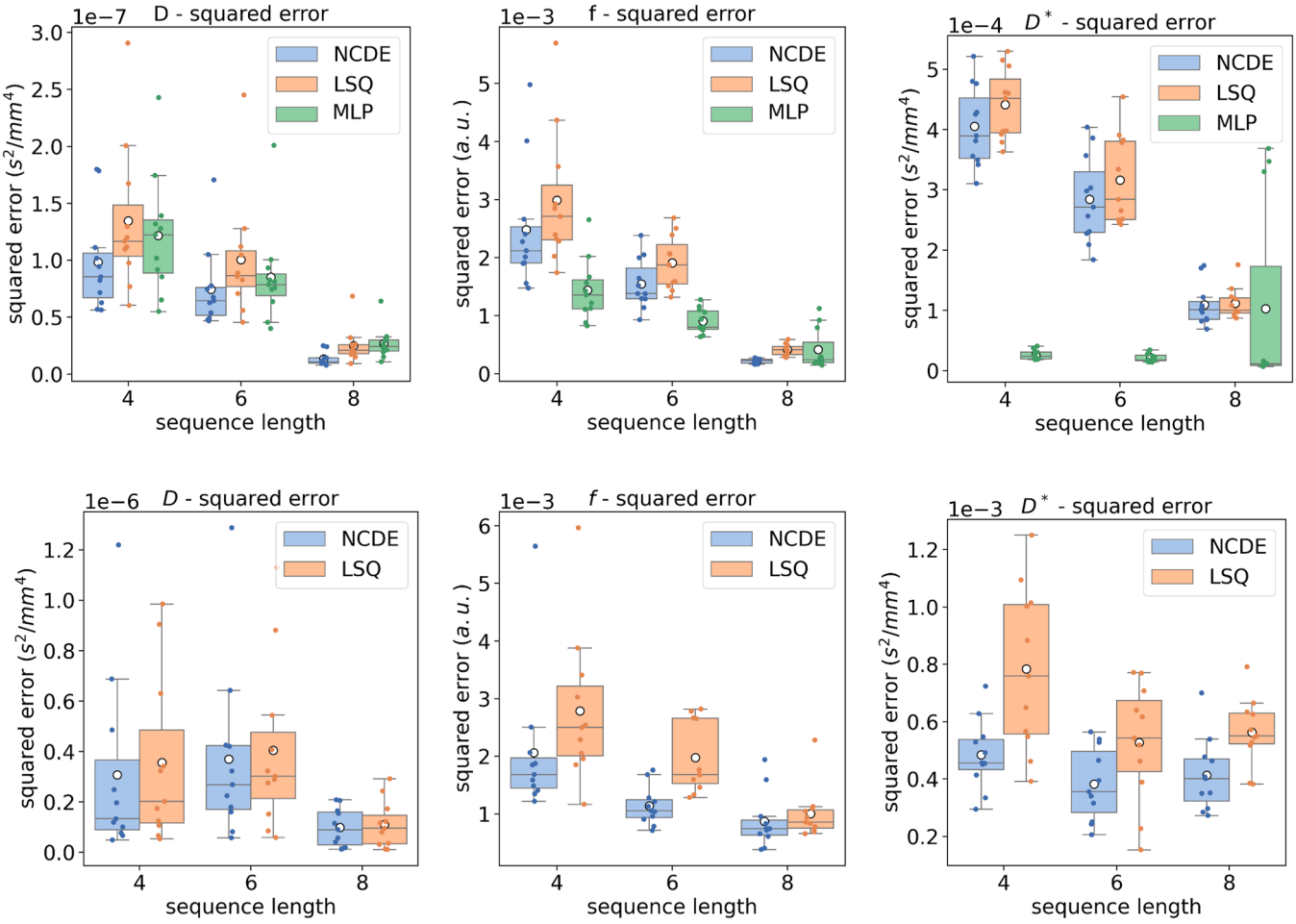

Fig. 8. Top row: box plot of mean squared errors per volunteer of the estimated IVIM parameters by NCDEs (blue), LSQ (orange) and MLP (green) as a function of sequence length. B-values used for the evaluation at 4, 6, 8 and 10 b-values were [0, 10, 50, 600] s/mm^2, [0, 10, 20, 50, 75, 600] s/mm^2, [0, 10, 20, 50, 75, 150, 400, 600] s/mm^2 and [0, 10, 20, 30, 50, 75, 150, 250, 400, 600] s/mm^2, respectively. Bottom row: box plot of mean squared errors per volunteer of the estimated IVIM parameter by NCDEs (blue), LSQ (orange) as a function of sequence length. Acquisition-specific MLPs cannot be evaluated at variable acquisition schemes. The white circle represents the mean squared error over all volunteers.

than LSQ fitting in abdominal and leg VFA *T1*-mapping but show smaller improvements in the brain, where higher SNR and higher image quality make QMRI parameters easier to estimate. The difference in relative performance across anatomical regions suggests that NCDE-based QMRI parameter estimates are more accurate than LSQ fitting under conditions of high uncertainty, similar to the behavior that is observed in simulations. In vivo, NCDEs yield more consistent IVIM MRI parameter estimates when the input sequence length is reduced compared to LSQ fitting. This work highlights the potential advantages of NCDE-based parameter estimation over LSQ fitting, particularly in challenging anatomical regions and high-uncertainty conditions, suggesting that NCDEs may offer a more robust approach for reliable QMRI.

With high SNR and densely sampled input sequences, the estimation of QMRI parameters is a trivial problem for which LSQ fitting performs sufficiently well. Unfortunately, such conditions are rare in clinical practice, which highlights that there is a need for suitable alternatives to LSQ fitting. Over the last 30 years, alternatives to LSQ fitting have been described in literature, including Bayesian inference (Barbieri et al., 2016; Baselice et al., 2016; Lofstedt et al., 2020; Neil et al., 1993; Scalco et al., 2022; Spinner et al., 2021; While, 2017) and neural networks (Bliesener et al., 2020; Epstein et al., 2024; Kaandorp et al., 2021; Ottens et al., 2022; Ulas et al., 2019; Vasylechko et al., 2022). The methods presented in these studies reduce the variation in estimation errors at low SNR levels, but are difficult to implement in a broader clinical context, for varying reasons. Bayesian inference is computationally intensive, time consuming and dependent on the choice of the prior distribution. Fully connected neural networks and convolutional neural networks are faster and learn prior distributions implicitly, but require a specifically trained neural network for each configuration of independent variables, creating considerable challenges for implementation and limiting reproducibility across different settings. Recurrent neural networks can handle variable length input sequences, but place assumptions on the regularity of the sampling interval which are often not met in quantitative MRI. Moreover, as training networks is a stochastic process, the need for retraining introduces an additional variation (Kaandorp et al., 2021) and hinders quality control. With the implementation of NCDEs, we aim to overcome most of these limitations. Our results suggest that NCDEs are able to reach similar accuracy as acquisition-specific MLPs, without the necessity to train and validate an MLP for every configuration of input variables. NCDEs are versatile networks that exhibit robustness to varied acquisition protocols, allowing end-users to directly apply a pre-trained, validated model, broadening access to QMRI parameter estimation for researchers and clinicians.

Although NCDEs have great potential, the NCDE-based parameter estimation method also has some specific requirements that may limit its applicability. For example, NCDE-based parameter estimation improved performance with input sequence normalization. In our implementation, this necessitated the acquisition of low flip angle images in VFA *T1*-mapping or images without diffusion weighting in IVIM MRI. For other implementations, other normalization techniques may be used. Further, we encountered difficulties when training the current NCDE implementation on longer sequences (>30 datapoints), leading to long training times and not finding adequate parameterizations of the NCDE in the training process. This is a known aspect of NCDEs (Kidger, 2022), with proposed solutions (Morrill, 2021; Walker, 2024) whose application in QMRI parameter estimation (e.g. densely sampled DCE-MRI) fall outside the scope of this work. Furthermore, since in our implementation the output of the NCDE and MLPs are rescaled to fall within a predefined range, the neural networks are specific to the predefined range of parameter values. Moreover, NCDEs currently only solve for 1 independent variable (e.g. varying b-value or varying flip angle). This means that more advanced models, with multiple varying acquisition parameters (e.g. VFA *T1*-mapping with varying repetition time (*TR*) and flip angle (*FA*) within one signal curve) cannot typically be solved. However, intelligent designs can help to generalize NCDE-based QMRI parameter estimation. For example for VFA *T1*-mapping, variations over multiple acquisition parameters among different acquisitions is solved by estimating $\frac{TR}{T1}$ instead, and multiplying the estimated $\frac{TR}{T1}$ with the a priori known $\frac{1}{TR}$. This allows the framework to generalize to acquisition series with any *TR*; however not to a series in which *TR* and *FA* are varied simultaneously. Note that varying TR and *FA* within the same experiment is not typically done in VFA *T1*-mapping. Finally, we demonstrated that NCDEs were faster in estimating parameters in VFA *T1*-mapping and DCE-MRI, but LSQ fitting only required CPUs while NCDEs required GPUs.

In this study, we demonstrate that NCDEs achieve performance comparable to acquisition-specific MLPs, both of which outperform LSQ fitting at realistic SNR levels. We are aware that MLPs do not necessarily represent the state of the art and that recent literature (Kaandorp et al., 2025; Ottens et al., 2022; Vasylechko et al., 2022) has shown that incorporating spatial awareness to networks for QMRI parameter estimation can lead to improved estimation accuracy. However, spatially-aware networks equally suffer from poor cross-protocol generalisability. We believe that solving the cross-protocol generalisability, potentially at the cost of a small reduction in performance as compared to spatially-aware networks, is a relevant hurdle. Potentially, combining NCDEs with spatial awareness can get best of both worlds, but is non-trivial and falls outside the scope of this research.

One limitation of our in vivo results is that the comparison between NCDEs, MLPs and LSQ fitting is constrained by the absence of definitive ground truth values. In VFA *T1*-mapping, we attempted to establish a robust reference standard using 30 flip angles, but nonetheless the resulting values remained susceptible to noise, image corruption, and fitting constraints. In IVIM MRI, it was not feasible to establish reliable reference values which were applicable to all estimation methods. The method-specific reference values that were used allowed us to assess consistency of the parameter estimates across different configurations of independent variables, but prevented us from drawing conclusions on the accuracy of the parameter estimates. Although this study serves as a proof-of-concept and is limited in terms of real-world generalizability, the results do indicate that NCDEs are capable of estimating QMRI parameters regardless of the acquisition scheme.

The field of deep learning-based QMRI parameter estimation is still in its early stages, leaving many questions unanswered. Although several studies have shown improvements of deep learning-based QMRI parameter estimation over the current state-of-the-art (Barbieri et al., 2020; Kaandorp et al., 2021; Ottens et al., 2022; Vasylechko et al., 2022), the underlying mechanisms for these improvements compared to LSQ fitting remain obscure. The limitations of the LSQ fitting algorithm are clear: it lacks prior information on expected distributions, does not enforce spatial coherence, is highly sensitive to outliers and noise and the complexity of QMRI forward models reduces the likelihood of reaching the global optimum in the optimization process. However, these very shortcomings contribute to its strength as LSQ fitting has become widely successful due to its simplicity, usability and predictability. NCDEs use a different approach, directly inferring the QMRI parameters from the signal curve without relying on an iterative process. This approach reduces their vulnerability to noise and the challenges posed by complex optimization landscapes. NCDEs have the potential to rival LSQ fitting in terms of usability and predictability, while offering advantages in finding QMRI parameter optima, ensuring spatial coherence and improving noise robustness and speed. Finally, the flexibility of NCDEs not only extends to a variety of MRI techniques, including pharmacokinetic DCE-MRI models, diffusion tensor imaging, and MR fingerprinting, but also to other modalities like positron emission tomography and ultrasound elastography, positioning NCDEs as versatile tools across imaging modalities.

In our eyes, the main advantage of NCDEs is that it allows training one network per QMRI technique that can be used around the world for any dataset, regardless of acquisition protocol or anatomical region. This

would have several advantages, including easy implementation, quality assurance of the model and comparable results across various sites. Our short-term goal is to train such models for IVIM MRI, DCE-MRI and VFA *T1*-mapping and share them with the community. Ultimately, to train such models with direct applicability in research and clinical settings, a large and varied dataset of quantitative data from multiple sites is required, with data from multiple vendors, field strengths, anatomical regions and acquisition protocols. To achieve this, we make an appeal to our readers. If the reader is willing to contribute to this dataset, we would appreciate if the reader could reach out to the authors.

5. Conclusion

NCDEs can function as a generic tool for accurate estimation of quantitative MRI parameters, irrespective of sequence length or quantitative MRI forward model. Overcoming these two hurdles opens up the use of neural networks for quantitative MRI parameter estimation to a broad audience of researchers and clinicians.

CRediT authorship contribution statement

Daan Kuppens: Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Data curation, Conceptualization. **Sebastiano Barbieri:** Writing – review & editing, Supervision, Methodology, Conceptualization. **Daisy van den Berg:** Software, Methodology. **Pepijn Schouten:** Writing – review & editing, Investigation. **Harriet C. Thoeny:** Writing – review & editing, Funding acquisition. **Hanneke W.M. van Laarhoven:** Writing – review & editing, Funding acquisition. **Myrte Wennen:** Writing – review & editing, Investigation. **Oliver J. Gurney-Champion:** Writing – review & editing, Supervision, Project administration, Methodology, Investigation, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Harriet C. Thoeny reports a relationship with Guerbet AG that includes: consulting or advisory. Hanneke W.M. van Laarhoven reports a consultant or advisory role for Amphera, Anocca, Astellas, AstraZeneca, BeiGene, Boehringer, BMS, Daiichi Sankyo, Dragonfly, MSD, Myeloid, and Servier; research funding, medication supply, or other research support from Auristone, Incyte, Merck, ORCA, and Servier; speaker roles for Astellas, AstraZeneca, BeiGene, Benecke, BMS, Daiichi Sankyo, JAAP, Medtalks, Novartis, Springer, and Travel Congress Management; and travel support from AstraZeneca. Other authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was funded by the KWF Dutch Cancer Society (KWF-UVA 2021.13785, OG-C and DK) and the Swiss National Science Foundation (32003B_176229/1, SB).

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.media.2025.103768.

References

Albers, G.W., 1998. Diffusion weighted MRI for evaluation of acute stroke. Neurology 51, S47–S49.

Barbieri, S., Donati, O.F., Froehlich, J.M., Thoeny, H.C., 2016. Impact of the calculation algorithm on biexponential fitting of diffusion-weighted MRI in upper abdominal organs. Magn. Reson. Med 75, 2175–2184. https://doi.org/10.1002/mrm.25765.

Barbieri, S., Gurney-Champion, O.J., Klaassen, R., Thoeny, H.C., 2020. Deep learning how to fit an intravoxel incoherent motion model to diffusion-weighted MRI. Magn. Reson. Med 83, 312–321. https://doi.org/10.1002/mrm.27910.

Baselice, F., Ferraioli, G., Pascazio, V., 2016. A Bayesian approach for relaxation times estimation in MRI. Magn. Reson. Imaging 34, 312–325. https://doi.org/10.1016/j.mri.2015.10.020.

Bliesener, Y., Acharya, J., Nayak, K.S., 2020. Efficient DCE-MRI parameter and uncertainty estimation using a neural network. IEEE. Trans. Med. Imaging 39, 1712–1723. https://doi.org/10.1109/TMI.2019.2953901.

Chauvie, S., Mazzoni, L.N., O'Doherty, J., 2023. A review on the use of imaging biomarkers in oncology clinical trials: quality assurance strategies for technical validation. Tomography 9, 1876–1902. https://doi.org/10.3390/tomography9050149.

Chen, R.T.Q., Rubanova, Y., Bettencourt, J., Duvenaud, D., 2018. Neural Ordinary Differential Equations. 32nd Conference on Neural Information Processing Systems (NeurIPS 2018). Montréal, Canada.

Christensen, K.A., Grant, D.M., Schulman, E.M., Walling, C., 1974. Optimal determination of relaxation times of fourier transform nuclear magnetic resonance. Determination of spin-lattice relaxation times in chemically polarized species. J. Phys. Chem. 78. https://doi.org/10.1021/j100612a022.

Epstein, S.C., Bray, T.J.P., Hall-Craggs, M., Zhang, H., 2024. Choice of training label matters: how to best use deep learning for quantitative MRI parameter estimation. Mach. Learn. Biomed. Imaging 2, 586–610. https://doi.org/10.59275/j.melba.2024-geb5.

Gupta, R.K., 1977. A new look at the method of variable nutation angle for the measurement of spin-lattice relaxation times using fourier transform NMR. J. Magn. Reson. 25 (1).

Gurney-Champion, O.J., Klaassen, R., Froeling, M., Barbieri, S., Stoker, J., Engelbrecht, M.R.W., Wilmink, J.W., Besselink, M.G., Bel, A., van Laarhoven, H.W.M., Nederveen, A.J., 2018. Comparison of six fit algorithms for the intra-voxel incoherent motion model of diffusion-weighted magnetic resonance imaging data of pancreatic cancer patients. PLoS. One 13. https://doi.org/10.1371/journal.pone.0194590.

Gurney-Champion, O.J., Landry, G., Redalen, K.R., Thorwarth, D., 2022. Potential of deep learning in quantitative Magnetic resonance imaging for personalized radiotherapy. Semin. Radiat. Oncol 32, 377–388. https://doi.org/10.1016/j.semradonc.2022.06.007.

Huang, H.M., 2022. An unsupervised convolutional neural network method for estimation of intravoxel incoherent motion parameters. Phys. Med. Biol 67. https://doi.org/10.1088/1361-6560/ac9a1f.

Kaandorp, M.P.T., Barbieri, S., Klaassen, R., van Laarhoven, H.W.M., Crezee, H., While, P.T., Nederveen, A.J., Gurney-Champion, O.J., 2021. Improved unsupervised physics-informed deep learning for intravoxel incoherent motion modeling and evaluation in pancreatic cancer patients. Magn. Reson. Med 86, 2250–2265. https://doi.org/10.1002/mrm.28852.

Kaandorp, M.P.T., Zijlstra, F., Karimi, D., Gholipour, A., While, P.T., 2025. Incorporating spatial information in deep learning parameter estimation with application to the intravoxel incoherent motion model in diffusion-weighted MRI. Med. Image. Anal 101, 103414. https://doi.org/10.1016/j.media.2024.103414.

Kety, S.S., 1951. The theory and applications of the exchange of inert gas at the lungs and tissues. Pharmocological. Rev. 3.

Khalifa, F., Soliman, A., El-Baz, A., Abou El-Ghar, M., El-Diasty, T., Gimel'farb, G., Ouseph, R., Dwyer, A.C., 2014. Models and methods for analyzing DCE-MRI: a review. Med Phys 41, 124301. https://doi.org/10.1118/1.4898202.

Kidger, P., Morrill, J., Foster, J., Lyons, T., 2020. In: Neural Controlled Differential Equations for Irregular Time Series, 34th. Vancouver, Canada.

Kidger, P., 2022. On neural differential equations (doctoral thesis). University of Oxford. https://doi.org/10.48550/arXiv.2202.02435.

Kingma, D.P., Ba, J.L., 2015. In: ADAM: A method for stochastic optimization. In: International Conference on Learning Representations (ICLR). San Diego, CA, USA.

Kurland, B.F., Gerstner, E.R., Mountz, J.M., Schwartz, L.H., Ryan, C.W., Graham, M.M., Buatti, J.M., Fennessy, F.M., Eikman, E.A., Kumar, V., Forster, K.M., Wahl, R.L., Lieberman, F.S., 2012. Promise and pitfalls of quantitative imaging in oncology clinical trials. Magn. Reson. Imaging 30, 1301–1312. https://doi.org/10.1016/j.mri.2012.06.009.

Liden, M., Adrian, D., Widell, J., Uggla, B., Thunberg, P., 2021. Quantitative T2* imaging of iron overload in a non-dedicated center - normal variation, repeatability and reader variation. Eur. J. Radiol. Open 8, 100357. https://doi.org/10.1016/j.ejro.2021.100357.

Ljimani, A., Caroli, A., Laustsen, C., Francis, S., Mendichovszky, I.A., Bane, O., Nery, F., Sharma, K., Pohlmann, A., Dekkers, I.A., Vallee, J.P., Derlin, K., Notohamiprodjo, M., Lim, R.P., Palmucci, S., Serai, S.D., Periquito, J., Wang, Z.J., Froeling, M., Thoeny, H.C., Prasad, P., Schneider, M., Niendorf, T., Pullens, P., Sourbron, S., Sigmund, E.E., 2020. Consensus-based technical recommendations for clinical translation of renal diffusion-weighted MRI. MAGMA 33, 177–195. https://doi.org/10.1007/s10334-019-00790-y.

Lofstedt, T., Hellstrom, M., Bylund, M., Garpebring, A., 2020. Bayesian non-linear regression with spatial priors for noise reduction and error estimation in quantitative MRI with an application in *T1* estimation. Phys. Med. Biol 65, 225036. https://doi.org/10.1088/1361-6560/abb9f5.

Manfrini, E., Smits, M., Thust, S., Geiger, S., Bendella, Z., Petr, J., Solymosi, L., Keil, V.C., 2021. From research to clinical practice: a European neuroradiological survey on quantitative advanced MRI implementation. Eur. Radiol 31, 6334–6341. https://doi.org/10.1007/s00330-020-07582-2.

Morrill, J., Salvi, C., Kidger, P., Foster, J., 2021. Neural rough differential equations for long time series. In: Proceedings of the 38th International Conference on Machine Learning. Virtual event, 7829-7838.

Neil, J.J., Bretthorst, G.L., 1993. On the use of bayesian probability theory for analysis of exponential decay data: an example taken from Intravoxel incoherent motion experiments. Magn. Reson. Med. 29, 642–647. https://doi.org/10.1002/mrm.1910290510.

O'Connor, J.P.B., Aboagye, E.O., Adams, J.E., Aerts, H.J.W.L., Barrington, S.F., Beer, A. J., Boellaard, R., Bohndiek, S.E., Brady, M., Brown, G., Buckley, D.L., Chenevert, T. L., Clarke, L.P., Collette, S., Cook, G.J., deSouza, N.M., Dickson, J.C., Dive, C., Evelhoch, J.L., Faivre-Finn, C., Gallagher, F.A., Gilbert, F.J., Gillies, R.J., Goh, V., Griffiths, J.R., Groves, A.M., Halligan, S., Harris, A.L., Hawkes, D.J., Hoekstra, O.S., Huang, E.P., Hutton, B.F., Jackson, E.F., Jayson, G.C., Jones, A., Koh, D.M., Lacombe, D., Lambin, P., Lassau, N., Leach, M.O., Lee, T.Y., Leen, E.L., Lewis, J.S., Liu, Y., Lythgoe, M.F., Manoharan, P., Maxwell, R.J., Miles, K.A., Morgan, B., Morris, S., Ng, T., Padhani, A.R., Parker, G.J.M., 2017. In: Pathak, A.P., Peet, A.C., Punwani, S., Reynolds, A.R., Robinson, S.P., Shankar, L.K., Sharma, R.A., Soloviev, D., Stroobants, S., Sullivan, D.C., Taylor, S.A., Tofts, P.S., Tozer, G.M., van Herk, M., Walker-Samuel, S., Wason, J., Williams, K.J., Workman, P., Yankeelov, T.E., Brindle, K.M., McShane, L.M., Jackson, A., Waterton, J.C., 2017. Imaging biomarker roadmap for cancer studies. Nat. Rev. Clin. Oncol. 14, 169–186.

Orton, M.R., d'Arcy, J.A., Walker-Samuel, S., Hawkes, D.J., Atkinson, D., Collins, D.J., Leach, M.O., 2008. Computationally efficient vascular input function models for quantitative kinetic modelling using DCE-MRI. Phys. Med. Biol 53, 1225–1239. https://doi.org/10.1088/0031-9155/53/5/005.

Ottens, T., Barbieri, S., Orton, M.R., Klaassen, R., van Laarhoven, H.W.M., Crezee, H., Nederveen, A.J., Zhen, X., Gurney-Champion, O.J., 2022. Deep learning DCE-MRI parameter estimation: application in pancreatic cancer. Med. Image. Anal 80, 102512. https://doi.org/10.1016/j.media.2022.102512.

Rata, M., Collins, D.J., Darcy, J., Messiou, C., Tunariu, N., Desouza, N., Young, H., Leach, M.O., Orton, M.R., 2015. Assessment of repeatability and treatment response in early phase clinical trials using DCE-MRI: comparison of parametric analysis using MR- and CT-derived arterial input functions. Eur. Radiol 26, 1991–1998. https://doi.org/10.1007/s00330-015-4012-9.

Rosenkrantz, A.B., Mendiratta-Lala, M., Bartholmai, B.J., Ganeshan, D., Abramson, R.G., Burton, K.R., Yu, J.P.J., Scalzetti, E.M., Yankeelov, T.E., Subramaniam, R.M., Lenchik, L., 2015. Clinical utility of quantitative imaging. Acad. Radiol 22, 33–49. https://doi.org/10.1016/j.acra.2014.08.011.

Scalco, E., Mastropietro, A., Rizzo, G., Lanzarone, E., 2022. A clustering approach to improve IntraVoxel incoherent motion maps from DW-MRI using conditional auto-regressive bayesian model. Appl. Sci. 12. https://doi.org/10.3390/app12041907.

Sourbron, S.P., Buckley, D.L., 2013. Classic models for dynamic contrast-enhanced MRI. NMR Biomed 26, 1004–1027. https://doi.org/10.1002/nbm.2940.

Spinner, G.R., Federau, C., Kozerke, S., 2021. Bayesian inference using hierarchical and spatial priors for intravoxel incoherent motion MR imaging in the brain: analysis of cancer and acute stroke. Med. Image. Anal 73, 102144. https://doi.org/10.1016/j.media.2021.102144.

Taylor, A.J., Salerno, M., Dharmakumar, R., Jerosch-Herold, M., 2016. T1 Mapping: basic techniques and clinical applications. JACC. Cardiovasc. Imaging 9, 67–81. https://doi.org/10.1016/j.jcmg.2015.11.005.

Tofts, P.S., Brix, G., Buckley, D.L., Evelhoch, J.L., Henderson, E., Knopp, M.V., Larsson, H.B.W., Lee, T.Y., Mayr, N.A., Parker, G.J.M., Port, R.E., Taylor, J., Weisskopf, R.M., 1999. Estimating kinetic parameters from dynamic contrast-enhanced *T1*-weighted MRI of a diffusable tracer: standardized quantities and symbols. J. Magn. Reson. Imaging 10, 223–232, 10.1002/(SICI)1522-2586(199909)10:3<223::AID-JMRI2>3.0.CO;2-S.

Ulas, C., Das, D., Thrippleton, M.J., Valdes Hernandez, M.D.C., Armitage, P.A., Makin, S. D., Wardlaw, J.M., Menze, B.H., 2019. Convolutional neural networks for direct inference of pharmacokinetic parameters: application to stroke dynamic contrast-enhanced MRI. Front. Neurol 9, 1147. https://doi.org/10.3389/fneur.2018.01147.

van Houdt, P.J., Ragunathan, S., Berks, M., Ahmed, Z., Kershaw, L.E., Gurney-Champion, O.J., Tadimalla, S., Arvidsson, J., Sun, Y., Kallehauge, J., Dickie, B., Levy, S., Bell, L., Sourbron, S., Thrippleton, M.J., 2023. Contrast-agent-based perfusion MRI code repository and testing framework: ISMRM Open Science Initiative for Perfusion Imaging (OSIPI). Magn. Reson. Med 91, 1774–1786. https://doi.org/10.1002/mrm.29826.

Vasylechko, S.D., Warfield, S.K., Afacan, O., Kurugol, S., 2022. Self-supervised IVIM DWI parameter estimation with a physics based forward model. Magn. Reson. Med 87, 904–914. https://doi.org/10.1002/mrm.28989.

Walker, B., McLeod, A.D., Qin, T., Cheng, Y., Li, H., Lyons, T., 2024. Log neural controlled differential equations: the lie brackets make a difference. In: Proceedings of the 41st International Conference on Machine Learning. Vienna, Austria.

While, P.T., 2017. A comparative simulation study of bayesian fitting approaches to intravoxel incoherent motion modeling in diffusion-weighted MRI. Magn. Reson. Med. 78, 2373–2387. https://doi.org/10.1002/mrm.26598.