



# Bayesian inference using hierarchical and spatial priors for intravoxel incoherent motion MR imaging in the brain: Analysis of cancer and acute stroke

Georg Ralph Spinner<sup>1</sup>, Christian Federau, Sebastian Kozerke\*

Institute for Biomedical Engineering, University and ETH Zurich, Gloriastrasse 35, Zurich 8092, Switzerland

## ARTICLE INFO

### Article history:

Received 21 November 2020

Revised 12 June 2021

Accepted 21 June 2021

Available online 29 June 2021

### Keywords:

Intravoxel incoherent motion imaging

Bayesian inference

Cancer

Acute stroke

## ABSTRACT

The intravoxel incoherent motion (IVIM) model allows to map diffusion ( $D$ ) and perfusion-related parameters ( $F$  and  $D^*$ ). Parameter estimation is, however, error-prone due to the non-linearity of the signal model, the limited signal-to-noise ratio (SNR) and the small volume fraction of perfusion in the *in-vivo* brain. In the present work, the performance of Bayesian inference was examined in the presence of brain pathologies characterized by hypo- and hyperperfusion. In particular, a hierarchical and a spatial prior were combined. Performance was compared relative to conventional segmented least squares regression, hierarchical prior only (non-segmented and segmented data likelihoods) and a deep learning approach. Realistic numerical brain IVIM simulations were conducted to assess errors relative to ground truth. *In-vivo*, data of 11 central nervous system cancer patients and 9 patients with acute stroke were acquired. The proposed method yielded reduced error in simulations for both the cancer and acute stroke scenarios compared to other methods across the whole investigated SNR range. The contrast-to-noise ratio of the proposed method was better or on par compared to the other techniques *in-vivo*. The proposed Bayesian approach hence improves IVIM parameter estimation in brain cancer and acute stroke.

© 2021 The Authors. Published by Elsevier B.V.

This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

## 1. Introduction

In the presence of diffusion encoding gradients, perfusion of tissue leads to a modulation of the MR signal. The influence of perfusion and water self-diffusion on the MR signal can be described by the intravoxel incoherent motion (IVIM) theory (Le Bihan et al., 1986; Le Bihan, 2019), which models perfusion as a pseudo-diffusion process on the macroscopic scale. Similar to water self-diffusion (coefficient  $D$ ), perfusion is described to cause a mono-exponential MR signal decay with increasing diffusion weighting. The perfusion fraction  $F$  depends on the vascular volume fraction in a given voxel relative to the water contained in the extravascular space. The pseudo-diffusion coefficient  $D^*$  describes blood flow through arterioles and capillaries as a diffusion-like process and is proportional to blood flow velocity and mean vessel segment length. Considering both diffusion and perfusion, the IVIM

model results, to first approximation, in a bi-exponential decay of the image magnitude as a function of diffusion weighting ( $b$ -value). The model has frequently been used to fit data acquired in a variety of anatomical regions (Lemke et al., 2009; Luciani et al., 2008; Notohamiprodjo et al., 2015; Wetscherek et al., 2015; Yamada et al., 1999), heart (Moulin et al., 2016; Spinner et al., 2017) and brain (Federau et al., 2012; Federau, 2017).

In the brain, IVIM has been used to investigate cancer (Bisdas et al., 2013; Federau et al., 2014a; 2014b; 2017; Togao et al., 2016). More specifically, IVIM was found to be a valuable technique to assess perfusion in central nervous cancer, where the perfusion fraction  $F$  (Federau et al., 2017) and  $D^*$  (Puig et al., 2016) were prognostic for survival.

In acute stroke assessment (Suo et al., 2016), IVIM benefits from its endogenous contrast mechanism, which renders it independent from the arterial input function or any dispersion effects of exogenous contrast-agent techniques. Accordingly, no vein puncture or contrast-agent administration is needed. It was shown that IVIM perfusion fraction was significantly reduced in the visible infarct (Federau et al., 2014c). Furthermore, IVIM was used to assess the quality of the collateral blood flow in hyper-acute stroke

\* Corresponding author.

E-mail address: [kozerke@biomed.ee.ethz.ch](mailto:kozerke@biomed.ee.ethz.ch) (S. Kozerke).

<sup>1</sup> Present address: Biomedical Simulation Lab, Zurich University of Applied Sciences (ZHAW), Schloss 1, 8820 Wädenswil, Switzerland

(Federau et al., 2019). Hence, IVIM is a useful tool to assess diffusion and perfusion in two frequent brain pathologies, namely cancer and acute stroke.

Various improvements to brain IVIM data processing have been proposed to increase parameter estimation quality. It was noted that a segmented approach of first determining the diffusion contribution prior to fitting the perfusion parameters was superior to a single-step fitting approach (Pekar et al., 1992). Parameter estimation quality of segmented non-linear least squares (LSQ) regression was assessed in simulations and in brains of healthy volunteers (Wu et al., 2015). The non-segmented approach was compared to the segmented model for non-linear LSQ fitting both in simulations and in brain cancer patients (Meeus et al., 2017). Non-linear segmented LSQ regression was further used to study hyper-oxygenation-induced vasoconstriction in the human brain (Federau et al., 2012) and glioma patients (Federau et al., 2014a). In addition, non-negative LSQ fitting was investigated and compared to both the segmented and the non-segmented approaches in malignant lesions (glioma, and metastasis), benign lesions (meningioma) and multiple sclerosis (Keil et al., 2017).

Besides the LSQ approaches, Bayesian inference of IVIM parameters has been used in the rat brain (Neil and Bretthorst, 1993), human brain (Rydhög et al., 2014), the liver (Dyvorne et al., 2013) and other abdominal organs (Barbieri et al., 2016). An analysis of various prior distributions has been conducted in a tumor mouse model (Gustafsson et al., 2018). A hierarchical Bayesian network, which imposes a global prior on all voxels of interest was proposed for data of the liver (Orton et al., 2014) and later applied to the heart (Spinner et al., 2017). This method was compared to flat priors in pancreatic patients (Gurney-Champion et al., 2018). In addition to these global approaches, which ignore spatial correlations, local spatial regularization has been exploited by using Markov random fields (MRF) in the liver (Freiman et al., 2013). This approach was compared to the hierarchical Bayesian approach in simulations and *in-vivo* data of the liver (While, 2017) and in breast cancer (Vidić et al., 2019). Recently, a deep learning approach was presented by Barbieri et al. (2020) and later modified by (Kaandorp et al., 2021).

Alongside advances in IVIM parameter inference,  $b$ -value selection (Lemke et al., 2011) during data acquisition has been optimized using Monte Carlo simulations. Besides improvements to the MR sequence design (Spinner et al., 2018), cerebral blood volume contrast and contrast-to-noise ratio (CNR) enhancements through  $T_2$ -preparation (Federau and O'Brien, 2015) have been achieved.

Despite all advances, reliable IVIM in the *in-vivo* brain remains challenging due to the inherently small perfusion volume fraction (Pekar et al., 1992) and the low signal-to-noise ratio (SNR) achievable *in-vivo*, due to time constraints of clinical workflows (Federau et al., 2014a; 2014c). The non-linearity of the IVIM model creates an ill-posed inversion problem of estimating the IVIM parameters. The resulting parameter estimation uncertainties frequently lead to erroneous IVIM parameter maps, where pathological lesions can be obscured by speckle-like noise and parameter contrast can be insufficient, especially for  $D^*$  (Federau et al., 2014a; 2014c).

The objective of the present work was to advance parameter inference using an improved Bayesian method and evaluate it in the presence of frequent pathologies, namely brain cancer and acute stroke. To this end, a novel Bayesian inference approach is presented, which combines and leverages the following aforementioned techniques together with a dedicated inference method:

- Spatial prior in the form of an edge-preserving Markov random field
- Hierarchical multivariate Normal prior for partial pooling of estimates

- Hyper-parameters of the hierarchical prior are pre-learned from data using Jeffreys prior and a segmented data likelihood
- Inference via Markov chain Monte Carlo using the Metropolis-Hastings algorithm for the combination of spatial and hierarchical priors

## 2. Theory

### 2.1. IVIM model

The IVIM model (Le Bihan et al., 1986) is a scalar two-compartment model of bi-exponential form with  $b$ -value ( $s/mm^2$ ) dependent magnitude  $S$  according to

$$S(b) = S_0[F \exp(-bD^*) + (1 - F) \exp(-bD)] \quad (1)$$

with diffusion coefficient  $D$  ( $mm^2/s$ ), perfusion fraction  $F$  (%) and pseudo-diffusion coefficient  $D^*$  ( $mm^2/s$ ). The magnitude without diffusion weighting is denoted  $S_0$ . Parameter estimation quality is typically improved by applying a segmented (segm.) approach (Federau et al., 2012; 2014a; Spinner et al., 2018), where only diffusion is expected to influence the recorded signal for  $b \geq b_s$  and  $b_s$  is the value which splits the  $b$ -value range such that:

$$S(b) \approx S_{int}^{Diff} \exp(-bD) \quad (2)$$

yielding  $D$  and the intercept  $S_{int}^{Diff} = S_0(1 - F)$ . Subsequently, the full IVIM model according to Eq. (1) is fitted to determine the remaining parameters with the two previously determined parameters fixed.

### 2.2. Data likelihood

For a SNR larger or equal 5, the noise distribution can be approximated by a Normal distribution (Gudbjartsson and Patz, 1995). Hence, a Normal data likelihood is used for the IVIM model in Eq. (1) and the diffusion-only model in Eq. (2) in the  $i$ th voxel ( $m$  in total):

$$p(\mathbf{y}_i | \theta_i) = \prod_{j=1}^n N(y_{i,j} | S(b_j, \theta_i), \sigma_{yi}^2) \quad (3)$$

where  $\mathbf{y}_i$  is a vector of length  $n$  stacking  $b$ -value dependent magnitude data per voxel. Unknown IVIM/diffusion parameters in voxel  $i$  are noted in shorthand here as  $\theta_i$ . The local image noise standard deviation is denoted  $\sigma_{yi}$ .

Nuisance parameters such as  $S_0$  and  $\sigma_y$  can be marginalized out in a Bayesian fashion (Orton et al., 2014). This can be achieved by choosing a prior distribution in the form of a product of a Normal and an inverse Gamma distribution for each voxel  $i$ :

$$p(S_{0i}, \sigma_{yi}^2) = N(S_{0i} | 0, \delta^2 \sigma_{yi}^2 / (\mathbf{g}_i^T \mathbf{g}_i)) \cdot IG(\sigma_{yi}^2 | \alpha, \beta) \quad (4)$$

with constant  $\delta^2$  as proposed in Zellners  $g$ -prior (De Finetti et al., 1986), inverse Gamma shape parameter  $\alpha$ , scale parameter  $\beta$  and  $\mathbf{g}_i$  being the vector modelling the fixed effects for all  $n$   $b$ -values, i.e. the IVIM model in Eq. (1) or the diffusion-only signal described in Eq. (2), both without proportionality constant  $S_0$ . This conjugate prior choice allows for an analytical evaluation. A lack of prior knowledge about the nuisance parameters can be addressed by taking the limits  $\delta \rightarrow \infty$  and  $\alpha, \beta \rightarrow 0$ , which correspond to infinitely broad prior distributions. The marginalized data likelihood at these limits is then given for the  $i$ th voxel as:

$$p(\mathbf{y}_i | \theta_i) \propto [\mathbf{y}_i^T \mathbf{y}_i - (\mathbf{y}_i^T \mathbf{g}_i)^2 / \mathbf{g}_i^T \mathbf{g}_i]^{-\frac{n}{2}} \quad (5)$$

For the segmented approach, the data likelihood can be decomposed into a product

$$p(\mathbf{y}_i | \theta_i) = \prod_{j=1}^{s-1} N(y_{i,j} | S(b_j, \theta_i), \sigma_{y_i}^2) \cdot \prod_{k=s}^n N(y_{i,k} | S(b_k, \theta_i), \sigma_{y_i}^2) \quad (6)$$

For  $b < b_s$  the IVIM model from Eq. (1) was considered for  $S(b_j, \theta_i)$ , while only diffusion was considered as fixed effect  $S(b_k, \theta_i)$  in voxel  $i$  as described in Eq. (2) for  $b \geq b_s$ . The marginalization procedure to remove the nuisance parameters as described above can then be performed in the same manner, but using the respective models  $g_{i,j}$  in the data likelihood in Eq. (5):

$$g_{i,j} = \begin{cases} F_i \exp(-b_j D_i^*) + (1 - F_i) \exp(-b_j D_i) & \text{if } b_j < b_s \\ (1 - F_i) \exp(-b_j D_i) & \text{if } b_j \geq b_s \end{cases} \quad (7)$$

### 2.3. Hierarchical Normal prior

In order to leverage information of all  $m$  voxels of interest, a hierarchical Bayesian network was assumed as data generating process. A (multivariate) Normal prior (Orton et al., 2014) was imposed on the log (D and D\*) and logit (F) transformed IVIM parameters  $\theta_i = [d_i, f_i, d_i^*] = [\log(D_i), \text{logit}(F_i), \log(D_i^*)]$  to allow for partial pooling of parameter estimates:

$$p(\theta_i | \mu, \Sigma_\mu) = |\Sigma_\mu|^{-\frac{1}{2}} \exp \left[ -\frac{1}{2} (\theta_i - \mu)^T \Sigma_\mu^{-1} (\theta_i - \mu) \right] \quad (8)$$

with hyper-parameters mean  $\mu$  and covariance matrix  $\Sigma_\mu$ .

### 2.4. Jeffreys hyper-prior

A hyper-prior was imposed on the hyper-parameters in a hierarchical fashion, yielding the so-called Bayesian shrinkage prior (BSP) method (Orton et al., 2014). Here, a data-driven prior in the form of Jeffreys prior (Jeffreys, 1946) was used, which is a so-called reference prior (Berger and Bernardo, 1992) and is defined for unknown mean and covariance of a multivariate Normal distribution as:

$$p(\mu, \Sigma_\mu) \propto \sqrt{|FI(\mu, \Sigma_\mu)|} = |\Sigma_\mu|^{-\frac{1}{2}} \quad (9)$$

where  $FI$  is the Fisher information matrix. The posterior distribution of the non-segmented BSP method as a product of data likelihood and prior(s) then reads:

$$\begin{aligned} p(\theta_{1...m}, \mu, \Sigma_\mu | \mathbf{y}_{1...m}) &\propto p(\mathbf{y}_{1...m} | \theta_{1...m}, \mu, \Sigma_\mu) \cdot p(\theta_{1...m}, \mu, \Sigma_\mu) \\ &= \prod_{i=1}^m \left[ p(\mathbf{y}_i | \theta_i) \cdot p(\theta_i | \mu, \Sigma_\mu) \right] \cdot p(\mu, \Sigma_\mu) \end{aligned} \quad (10)$$

The posterior for a segmented data likelihood accordingly reads (BSP segm.):

$$\begin{aligned} p(\theta_{1...m}, \mu, \Sigma_\mu | \mathbf{y}_{1...m}) &\propto \\ \prod_{i=1}^m \left[ p(\mathbf{y}_i^{1...s-1} | \theta_i) \cdot p(\mathbf{y}_i^{s...n} | \theta_i) \cdot p(\theta_i | \mu, \Sigma_\mu) \right] &\cdot p(\mu, \Sigma_\mu) \end{aligned} \quad (11)$$

Both the model terms  $\mathbf{g}$  and data vectors  $\mathbf{y}$  are segmented in Eq. (11):  $\mathbf{g}$  is defined in Eq. (7) and  $\mathbf{y}_i^{1...s-1}$  contains the data corresponding to  $0 \leq b \leq b_{s-1}$  and  $\mathbf{y}_i^{s...n}$  the remainder for every  $i$ th

voxel. For the original, non-segmented BSP method (Orton et al., 2014) in Eq. (10), no segmentation is used for the data  $\mathbf{y}$  and only the IVIM model is used in  $\mathbf{g}$ .

### 2.5. Spatial prior: Markov random field

In order to leverage local IVIM parameter similarity due to tissue homogeneity, a spatial prior in the form of a MRF was used (Freiman et al., 2013; Scannell et al., 2020). Accordingly, a generalised Gaussian MRF prior with zero mean was put on the differences between parameters of neighboring voxels:

$$p(\theta_i | \theta_j, v_{i,j}) \propto \exp \left[ -\frac{v_{i,j}}{2} \|W(\theta_i - \theta_j)\|_a^a \right] \quad (12)$$

where voxel  $j$  is a neighboring voxel of voxel  $i$ ,  $v_{i,j}$  is a rate parameter,  $W$  are weights and  $1 \leq a \leq 2$ . In order to preserve edges in the parameter maps,  $a = 1$  was chosen (Bardsley, 2012). Also,  $v_{i,j}$  was set to 1. The spatial prior then reads:

$$p(\theta_i | \theta_{n(i)}) \propto \prod_{p=1}^3 \exp \left[ -\frac{1}{2} \sum_{j \in n(i)} (W^p \cdot |\theta_i^p - \theta_j^p|) \right] \quad (13)$$

where  $n(i)$  are the neighboring voxels of voxel  $i$ ,  $\theta_i^p$  and  $W_i^p$  are the  $p$ th IVIM parameter and the corresponding  $p$ th weight in voxel  $i$ , respectively. The neighborhood size used in this study was  $5 \times 5$ . Voxels on the edge frequently had fewer neighboring voxels, because only the ones in the region-of-interest (ROI) were considered. Conditioned on fixed, i.e. data-derived estimates of the hyper-parameters  $\hat{\mu}$  and  $\hat{\Sigma}_\mu$ , the posterior including the spatial MRF prior reads:

$$p(\theta_{1...m} | \mathbf{y}_{1...m}) \propto \prod_{i=1}^m p(\mathbf{y}_i | \theta_i) \cdot p(\theta_i | \hat{\mu}, \hat{\Sigma}_\mu) \cdot p(\theta_i | \theta_{n(i)}) \quad (14)$$

The posterior in Eq. (14) hence combines a non-segmented data likelihood with a data-derived hierarchical prior as in the BSP method and a spatial MRF prior. Thus, this proposed method is termed BSP & MRF.

### 2.6. Markov chain Monte Carlo

In order to perform parameter inference under a given posterior, Markov chain Monte Carlo (MCMC) sampling can be used. Even though the integration to calculate expectation values under the posterior is analytically intractable, it is possible to draw samples and approximate the integrals with a sum in a Monte Carlo approach. In particular, it is possible to construct a Markov chain with a stationary distribution equal to the posterior distribution. Frequently used algorithms to generate the samples are e.g. the Metropolis-Hastings algorithm (Metropolis et al., 1953) and Gibbs sampling (Geman and Geman, 1984). For the posterior in Eq. (14) (BSP & MRF), the  $p$ th log/logit transformed IVIM parameter  $\hat{\theta}_i^p$  (i.e. estimates of  $d$ ,  $f$  and  $d^*$ ) in voxel  $i$  can be inferred:

$$\hat{\theta}_i^p = \int p(\theta_{1...m} | \mathbf{y}_{1...m}) \theta_i^p d\theta_{1...m} \approx \frac{1}{N} \sum_{j=1}^N \theta_i^{p,j} \quad (15)$$

where  $\theta_i^{p,j}$  is the  $j$ th sample of the  $p$ th log/logit transformed IVIM parameter and  $N$  is the amount of samples. An analogous approach for the posteriors in Eqs. (10) (BSP non-segm.) and (11) (BSP segm.) can be used, where also  $\mu$  and  $\Sigma_\mu$  can be inferred besides the voxel-wise log/logit transformed IVIM estimates. In order to assess convergence, multiple chains are used and the samples from all chains are combined. Note, that besides the mean as above, the median of the samples can also be used. The standard deviation of the samples can be furthermore used to assess the posterior width and hence estimation uncertainty, which is an intrinsic advantage of Bayesian methods.

### 3. Materials and methods

#### 3.1. Least squares fitting

The non-linear LSQ estimation of all IVIM parameters in Eq. (1) was implemented using Matlab (R2017b, Mathworks, USA) and the built-in function `lsqcurvefit`. Diffusivity larger than the self-diffusion of free water and pseudo-diffusivity was limited similar to Federau et al. (2012), while the perfusion fraction was not restricted:

$$\begin{aligned} 0 &\leq D \leq 2.5 \cdot 10^{-3} \text{ mm}^2/\text{s} \\ 0 &\leq F \leq 100\% \\ 0 &\leq D^* \leq 50 \cdot 10^{-3} \text{ mm}^2/\text{s} \end{aligned} \quad (16)$$

Both a non-segmented and a segmented fit were used according to Eqs. (1) and (2). For the segmented LSQ approach, a frequently used split  $b$ -value of  $b_s=200$  s/mm<sup>2</sup> was applied (Federau et al., 2014a; 2014c).

#### 3.2. Deep neural networks

IVIM fits using an optimized deep neural network (Kaandorp et al., 2021) were performed with publicly available code from <https://github.com/oliverchampion/IVIMNET> commit 56a9f08 using Python 3.6.12 and PyTorch 1.8.1. This method corresponded to the IVIM-NET<sub>optim</sub> technique as presented in Kaandorp et al. (2021) and was termed IVIMNET in this study for the sake of brevity. The public demo code was adapted to read-in brain data of this study and the parameter constraints were set to the ones in Eq. (16). Due to the stretching of the sigmoid activation functions beyond these constraints in IVIMNET, resulting negative parameter estimates were subsequently set to zero. The amount of data used for validation was reduced from 32 times the training batch size to 4 times the training batch size, because fewer voxels were available in this study compared to the data provided together with the demo code in the repository. Data was normalized by the  $b=0$  s/mm<sup>2</sup> images and the default constraints were used for  $S_0$  (0.7 and 2.0, respectively). All other settings of IVIMNET were left unchanged, apart from the modifications stated above. As proposed in Barbieri et al. (2020), Kaandorp et al. (2021), the neural networks were trained on the individual datasets and parameters were inferred subsequently on the same datasets.

#### 3.3. Bayesian inference

For all Bayesian methods, 4 Markov chains were started in parallel to speed up computation and assess convergence. The Metropolis-Hastings rate of acceptance was set to 0.234 (Gelman et al., 1997) and the proposal widths were adapted every 50 samples during the burn-in phase. The effectively used samples (after burn-in) from the Markov chains were pooled together and their median was used for parameter estimates. Parameter estimation uncertainty was assessed using the coefficient of variation of the pooled samples, i.e. the ratio of sample standard deviation and sample mean value of the IVIM estimates  $D$ ,  $F$  and  $D^*$ . The voxel-wise starting values were  $1 \cdot 10^{-3}$  mm<sup>2</sup>/s, 10% and  $10 \cdot 10^{-3}$  mm<sup>2</sup>/s with 10% Gaussian noise for  $D$ ,  $F$  and  $D^*$  respectively. These start values were used for all Bayesian inference procedures, apart from the proposed method (BSP & MRF), see below for details. If Jeffreys prior was used (BSP non-segm. & segm.),  $\mu$  and  $\Sigma_\mu$  were initialized with the mean and co-variance of the start values. Convergence was assessed by the  $\hat{R}$  statistic (Andrew Gelman and Donald B. Rubin, 1992) and a value of  $\hat{R} \leq 1.1$  (Kass et al., 1998) was considered satisfactory.

In order to assess LSQ estimation uncertainty, a non-segmented (Eq. (3)) and a segmented (Eq. (6)) version of the marginalized data

likelihood in Eq. (5) was employed together with uniform priors corresponding to a maximum likelihood approach with constraints as specified in Eq. (16). Parameters were hence not transformed (log/logit) for modelling and sampling as in the other Bayesian approaches. A burn-in phase of 1000 discarded samples was followed by another 1000 samples per Markov chain, yielding 4000 usable samples in total.

The hierarchical non-segmented BSP model was implemented as described originally (Orton et al., 2014) using a combination of Gibbs sampling and Metropolis-Hastings updates. IVIM parameters were accordingly log/logit transformed and the data likelihood was not segmented. A burn-in phase of 5000 discarded samples was followed by 5000 samples per Markov chain (in total 20,000 samples used) to account for the relatively slow convergence of the hyper-parameters  $\mu$  and  $\Sigma_\mu$ . A modification (BSP segm.) with a segmented data likelihood and a split  $b$ -value of  $b_s=800$  s/mm<sup>2</sup> was also implemented and used with otherwise same settings. A high split  $b$ -value was chosen to reduce perfusion influence on the high  $b$ -value magnitude and hence parameter bias, after checking results from a parameter sweep of  $b_s$  in simulated data.

The proposed method (BSP & MRF) uses a non-segmented data likelihood and combines a hierarchical (BSP) with a spatial prior (MRF). The fixed hyper-parameters  $\hat{\mu}$  and  $\hat{\Sigma}_\mu$  were learnt from data using estimates from a previous segmented BSP approach, as explained above. The weights  $W$  in the spatial prior were set to the inverse of the current voxel-wise parameter estimate in the Metropolis-Hastings algorithm (Scannell et al., 2020). This normalization accounts for the different scaling of the parameters. To speed up computation of the relatively expensive MRF computation, the estimates from the segmented BSP approach were used as start values. A burn-in phase of 1000 discarded samples was followed by 1000 samples per Markov chain, yielding 4000 usable samples in total. Results from a single Markov chain with the same burn-in length, but of four-fold length were assessed for the proposed method in simulations with realistic SNR.

All calculations were performed on a Linux workstation with Intel Xeon E5 CPU with 2.5 GHz and 128 GB RAM, except for the IVIMNET calculations which were performed on a Windows laptop with Intel i5 CPU with 2.9 GHz and 16 GB RAM.

All presented Bayesian inference methods were implemented in Matlab (R2017b, Mathworks, USA) and code is available under: <https://github.com/georgspinner/BayesIVIM>.

#### 3.4. Simulations

MR data generation in a realistic numerical brain phantom was performed using a modified version of existing brain simulation routines (Guerquin-Kern et al., 2012) with eight surface coils. Cancer-like lesions were drawn in the frontal region using Bézier-curves consisting of two outer layers (cancer and edema) with increased diffusion and perfusion, while the central necrotic region had (virtually) zero perfusion. In a second experiment, an acute stroke-like lesion was modeled by a core with (virtually) zero perfusion surrounded by a penumbra of reduced perfusion. Diffusion was reduced across the whole lesion.

The non-diffusion weighted magnitude  $S_0$  was set to higher values compared to the healthy parenchyma in both scenarios to reflect the situations *in-vivo*. Data was weighted with the IVIM model magnitude of Eq. (1) using IVIM parameters as noted in Table 1. Simulated  $b$ -values were: 0, 10, 20, 40, 80, 110, 140, 170, 200, 300, 400, 500, 600, 700, 800 and 900 s/mm<sup>2</sup>. Normal distributed noise was added to the real and imaginary parts of the signal such that the mean SNR across the ROI equaled 10, 20, 30, 40, 50, 75, 100, 125, 150, 175 and 200 after Roemer coil combination (Roemer et al., 1990). The simulation procedure is illustrated in Supplemental Fig. S1.



**Table 1**

Mean  $\pm$  standard deviation of IVIM parameter estimates from simulations. The asterisk (\*) indicates that estimates were significantly different from reference parameters ( $p < 0.05$ , Wilcoxon signed rank test with 5% significance level and Bonferroni correction for multiple testing). LSQ=least squares, IVIMNET=deep neural network, BSP=Bayesian shrinkage prior, MRF=Markov random field, segm.=segmented, WM=white matter, GM=gray matter.

		Cancer (SNR=40)					Acute stroke (SNR=30)			
		Necrotic core	WM	GM	Edema	Tumor	Necrotic core	Penumbra	WM	GM
D[10 <sup>-3</sup> mm <sup>2</sup> /s]	LSQ non-segm.	0.37 $\pm$ 0.25*	0.52 $\pm$ 0.21*	0.71 $\pm$ 0.25*	1.15 $\pm$ 0.19*	1.35 $\pm$ 0.15*	0.24 $\pm$ 0.10*	0.33 $\pm$ 0.16*	0.51 $\pm$ 0.23*	0.67 $\pm$ 0.29*
	LSQ segm.	0.38 $\pm$ 0.14	0.63 $\pm$ 0.11*	0.84 $\pm$ 0.07*	1.23 $\pm$ 0.05*	1.43 $\pm$ 0.06*	0.30 $\pm$ 0.06	0.42 $\pm$ 0.04*	0.64 $\pm$ 0.14*	0.84 $\pm$ 0.09*
	IVIMNET non-segm.	0.44 $\pm$ 0.07*	0.65 $\pm$ 0.05*	0.83 $\pm$ 0.04*	1.25 $\pm$ 0.03*	1.50 $\pm$ 0.03*	0.33 $\pm$ 0.03*	0.45 $\pm$ 0.02*	0.63 $\pm$ 0.08*	0.86 $\pm$ 0.04*
	BSP non-segm.	0.45 $\pm$ 0.05*	0.61 $\pm$ 0.04*	0.80 $\pm$ 0.03*	1.16 $\pm$ 0.03*	1.36 $\pm$ 0.03*	0.33 $\pm$ 0.02*	0.40 $\pm$ 0.02	0.57 $\pm$ 0.05*	0.79 $\pm$ 0.05*
	BSP segm.	0.45 $\pm$ 0.05*	0.61 $\pm$ 0.04*	0.80 $\pm$ 0.03*	1.16 $\pm$ 0.03*	1.37 $\pm$ 0.03*	0.31 $\pm$ 0.03*	0.40 $\pm$ 0.02	0.60 $\pm$ 0.05*	0.80 $\pm$ 0.04*
	BSP & MRF non-segm.	0.49 $\pm$ 0.05*	0.61 $\pm$ 0.03*	0.81 $\pm$ 0.03*	1.17 $\pm$ 0.03*	1.38 $\pm$ 0.03*	0.32 $\pm$ 0.02*	0.40 $\pm$ 0.02*	0.60 $\pm$ 0.04*	0.79 $\pm$ 0.04*
	Reference	0.40	0.60	0.80	1.20	1.40	0.30	0.40	0.60	0.80
F[%]	LSQ non-segm.	8.4 $\pm$ 12.3*	12.9 $\pm$ 16.9*	16.4 $\pm$ 19.2*	13.6 $\pm$ 11.7*	17.7 $\pm$ 8.0*	6.2 $\pm$ 10.5*	12.5 $\pm$ 19.0*	13.7 $\pm$ 17.6*	19.2 $\pm$ 22.2*
	LSQ segm.	3.9 $\pm$ 5.4	4.1 $\pm$ 4.6*	5.4 $\pm$ 3.2*	8.3 $\pm$ 2.2*	13.4 $\pm$ 2.8*	1.7 $\pm$ 2.3*	2.1 $\pm$ 2.0*	4.7 $\pm$ 5.5*	5.8 $\pm$ 4.1*
	IVIMNET non-segm.	0.1 $\pm$ 0.6*	2.6 $\pm$ 2.7*	7.1 $\pm$ 2.2*	1.0 $\pm$ 0.9*	0.0 $\pm$ 0.0*	0.1 $\pm$ 0.4*	0.9 $\pm$ 1.3*	3.9 $\pm$ 4.0*	4.8 $\pm$ 2.1*
	BSP non-segm.	3.1 $\pm$ 0.5*	4.7 $\pm$ 0.5*	7.2 $\pm$ 0.5*	11.8 $\pm$ 0.9*	16.0 $\pm$ 1.2*	26.6 $\pm$ 2.8*	18.0 $\pm$ 1.4*	9.9 $\pm$ 1.2*	7.4 $\pm$ 0.9*
	BSP segm.	3.1 $\pm$ 0.5*	4.7 $\pm$ 0.5*	7.1 $\pm$ 0.4*	11.5 $\pm$ 0.7*	15.5 $\pm$ 1.0*	3.0 $\pm$ 0.2*	3.8 $\pm$ 0.2*	5.5 $\pm$ 0.4*	7.2 $\pm$ 0.5*
	BSP & MRF non-segm.	3.8 $\pm$ 0.4*	4.9 $\pm$ 0.3*	6.8 $\pm$ 0.4*	11.1 $\pm$ 0.6*	14.6 $\pm$ 0.7*	3.4 $\pm$ 0.3*	3.9 $\pm$ 0.2*	5.7 $\pm$ 0.2*	6.9 $\pm$ 0.3*
	Reference	1.0	5.0	8.0	10.0	15.0	1.0	3.0	5.0	8.0
D*[10 <sup>-3</sup> mm <sup>2</sup> /s]	LSQ non-segm.	8.6 $\pm$ 15.0*	10.5 $\pm$ 15.5*	11.0 $\pm$ 13.3*	13.2 $\pm$ 10.4*	12.6 $\pm$ 6.0	7.7 $\pm$ 14.6*	7.9 $\pm$ 14.2	11.6 $\pm$ 16.6*	11.3 $\pm$ 14.1*
	LSQ segm.	31.6 $\pm$ 21.7*	24.7 $\pm$ 20.7*	17.7 $\pm$ 15.2*	15.9 $\pm$ 8.3*	15.7 $\pm$ 5.8*	29.7 $\pm$ 21.4*	25.1 $\pm$ 21.2*	26.4 $\pm$ 21.0*	19.6 $\pm$ 17.1*
	IVIMNET non-segm.	52.4 $\pm$ 8.6*	29.8 $\pm$ 8.9*	7.7 $\pm$ 2.1*	2.0 $\pm$ 0.2*	1.5 $\pm$ 0.2*	30.2 $\pm$ 2.5*	27.9 $\pm$ 2.0*	25.7 $\pm$ 5.4*	20.8 $\pm$ 3.3*
	BSP non-segm.	3.8 $\pm$ 0.4*	5.0 $\pm$ 0.3*	6.4 $\pm$ 0.3*	8.7 $\pm$ 0.8*	10.3 $\pm$ 0.8*	0.3 $\pm$ 0.1*	0.7 $\pm$ 0.1*	2.6 $\pm$ 0.8*	7.0 $\pm$ 1.3*
	BSP segm.	3.8 $\pm$ 0.4*	5.0 $\pm$ 0.3*	6.4 $\pm$ 0.3*	8.9 $\pm$ 0.5*	10.6 $\pm$ 0.6*	1.7 $\pm$ 0.2*	2.4 $\pm$ 0.2*	4.2 $\pm$ 0.5*	6.5 $\pm$ 0.5*
	BSP & MRF non-segm.	4.2 $\pm$ 0.3*	5.1 $\pm$ 0.2*	6.5 $\pm$ 0.3*	9.0 $\pm$ 0.5*	10.8 $\pm$ 0.5*	1.8 $\pm$ 0.2*	2.4 $\pm$ 0.1*	4.4 $\pm$ 0.3*	6.2 $\pm$ 0.4*
	Reference	0.1	4.0	6.0	10.0	12.0	0.1	3.0	4.0	6.0

### 3.5. Patient demographics

The cancer study was approved by the local ethics committee of the University of Lausanne. Patients were scanned from March 2011 to December 2012 and were also used as part of a clinical study cohort (Federau et al., 2014a). Consecutive histopathologic diagnostic findings were: 9 cancer patients had glioblastoma (WHO grade 4), 1 patient had oligoastrocytoma (WHO grade 3) and 1 patient had a neuroglial tumor (WHO grade 2). These patients had no relevant treatment history at the time of imaging (such as radio-, chemo- or antiangiogenic therapy). Out of the 11 cancer patients, 3 were female. The mean age of the cancer patients was 58 $\pm$ 15 years (age-range 24–84 years).

The two-center acute stroke study was approved by both local ethics committees at the Universities of Virginia and Lausanne. Data was collected from February 2011 to August 2013 in patients presenting with symptoms of hemispheric acute stroke and their data was also used as part of a clinical study cohort (Federau et al., 2014c). The mean time from symptom onset to imaging was 59 $\pm$ 46 h (range 10–120 h). Out of the 9 patients, 2 were female. The mean age of the stroke patients was 58 $\pm$ 20 years (age-range 25–86 years).

### 3.6. In-vivo data acquisition

The *in-vivo* acquisitions were performed using 3T MR scanners (Trio, Verio and Skyra; Siemens, Erlangen, Germany) with 32 multi-channel receive head coils. The imaging parameters were: slice thickness 4 mm, field-of-view of 297x297 mm<sup>2</sup> (cancer) and 270x270 mm<sup>2</sup> (acute stroke), spatial resolution=1.2x1.2 mm<sup>2</sup>, TR=4000 ms, TE=99 ms (cancer) and 89–102 ms depending on scanner (acute stroke), receiver bandwidth 1086 Hz/pixel (cancer) and 1106 Hz/pixel (acute stroke), spectrally selective fat saturation, parallel imaging with acceleration factor 2 and 75% partial Fourier in phase-encoding direction. Single-shot echo-planar readout was used for Stejskal and Tanner (1965) diffusion imaging with 16 *b*-values (same values as used for simulations) along 3 orthogonal directions, from which the trace was calculated by taking the mean.

A single average was acquired in an acquisition time of about 3 min.

### 3.7. Registration, segmentation & processing

IVIM parameters of the simulated data were estimated in all voxels of the brain, while excluding areas of simulated cerebrospinal fluid (CSF), bone and skin. Bias and standard deviation were assessed by calculating the absolute differences from the simulated parameters and the standard deviation of the estimated parameters within the different simulated tissues, respectively. The errors (bias and standard deviation) in different tissue sub-types (i.e. necrotic core, cancer and edema in the cancer scenario; necrotic core and penumbra in the stroke scenario; white and gray matter in healthy parenchyma) were then added as weighted sums for both the lesion and parenchyma, respectively. The weights of the sums accounted for the different numbers of simulated voxels of the different tissue types.

Images of the *in-vivo* data were first registered slice-wise after magnitude normalization and taking the square-root. To this end, a nuclear (i.e. principal component analysis) group-wise metric was used to measure image dissimilarity during registration together with isotropic total variation regularization of displacements (Vishnevskiy et al., 2017) using code from <https://github.com/visva89/pTVreg>.

The *in-vivo* data was then automatically segmented in a heuristic approach. First the foreground was selected by magnitude thresholding with lower (15% of maximum signal) and upper (40% of minimal signal) limits, followed by global image thresholding using Otsu's method (Otsu, 1979) and subsequent erosion and dilation of the mask (disk-shaped structure elements with radius 3 and 10, respectively). The pre-segmentation was used as starting point for active contours (Chan and Vese, 2001) segmentation with 100 iterations. The resulting mask was opened (disk-shaped structure element with radius 9). In an additional step, voxels with  $D \geq 1.25 \cdot 10^{-3}$  mm<sup>2</sup> and  $F \geq 25\%$  (fitted using a log-linear LSQ fit of data with  $b \geq 200$  s/mm<sup>2</sup>) were excluded, to prevent inclusion of CSF. Finally, the resulting mask was eroded and closed (disk-

shaped structure element with radius 1) to remove detached voxels around the edges of the ROI.

From the *in-vivo* data, a single slice was chosen in each patient and ROIs were placed manually in consensus of two experienced neuroradiologists in the cancer patients (Federau et al., 2014a), with high IVIM perfusion fraction (cancer lesion) and on the contralateral white matter parenchyma (control). The IVIM parameters for the segmentation were estimated using segmented non-linear LSQ with a split  $b$ -value of 200 s/mm<sup>2</sup>. Cystic, hemorrhagic, or necrotic areas were avoided. For acute stroke patients, ROIs in the ischemic core were obtained by thresholding the diffusion map (from segmented non-linear LSQ) on a single axial slice with the largest area of infarction. The ROIs were then verified by an experienced neuroradiologist and manually corrected if necessary (Federau et al., 2014c). All ROIs were placed such that they included as little CSF or as few large vessels as possible. The manually drawn ROIs in the lesion and control parenchyma were finally added to the segmentation (if not already automatically included).

*In-vivo*SNR was estimated using the dual acquisition method (Dietrich et al., 2007; Reeder et al., 2005), where the three diffusion direction weightings with  $b=10$  s/mm<sup>2</sup> served as signal averages. The mean SNR of the three acquisition pairs was calculated for each brain segmentation in every patient. Finally, the reported SNR of the trace takes into account the number of diffusion gradient directions, which were averaged to yield trace magnitude images. For both *in-vivo* scenarios, the SNR was reported as mean±standard deviation across all segmented voxels of all data sets.

CNRs were obtained for each IVIM parameter by subtracting the mean IVIM parameter estimate in the lesion from the mean estimate in the control region and dividing it by the standard deviation of the estimates within the control region (Federau and O'Brien, 2015). For acute stroke data, the subtraction was vice versa, such that CNR was expected to be positive in the respective lesions for both cancer and acute stroke. For both *in-vivo* scenarios, the CNR was reported as median±inter-quartile range across all data sets.

### 3.8. Statistical analysis

Differences of the IVIM estimates from the simulated parameters were assessed using a Wilcoxon signed-rank test with 5% significance level with Bonferroni correction for multiple testing. Differences of the IVIM estimates *in-vivo* between lesion and control parenchyma were tested by using Wilcoxon rank sum/Mann-Whitney  $U$  test (5% significance level with Bonferroni correction for multiple testing). Differences in CNR *in-vivo* among the regression methods were tested using the Kruskal-Wallis test (5% significance level) with a subsequent multiple comparison test for unequal medians.

## 4. Results

### 4.1. Simulations

The estimated IVIM parameters are presented together with the simulated parameters in Table 1. The proposed method (BSP & MRF) produced accurate estimates for both scenarios of cancer (SNR=40) and acute stroke (SNR=30) with a relative error (bias) below 30% for all IVIM parameters, apart from the necrotic cores. There, the perfusion parameters (F and D\*) are over-estimated, but they remained smaller than the estimates from the segmented LSQ fit. The corresponding standard deviations were the lowest of all methods, especially considering F and D\*. The LSQ methods yielded large standard deviations in both scenarios. The IVIM-NET methods produced estimates, which were generally compara-

ble to LSQ and Bayesian methods. However, perfusion was estimated to be virtually zero in the cancer lesion and D\* was over-estimated in healthy parenchyma for both simulated pathologies. The non-segmented BSP method produced accurate estimates for D in both scenarios, but over-estimated F and under-estimated D\* in most tissue types in the acute stroke scenario. Computation times for parameter estimation in a single simulation dataset were about 5/10/15/45 min for LSQ/IVIMNET/BSP methods/BSP & MRF on 1/1/4/4 CPU cores. The mean differences of the parameter estimates in cancer (SNR=40) and acute stroke (SNR=30) from multiple Markov chains compared to a single longer Markov chain were found to be smaller than the reported parameter accuracy.

Results for the proposed method (BSP & MRF) from a parameter sweep of the splitting  $b$ -value  $b_s$  can be found in the Supplemental Fig. S2. It was generally found that parameter estimation error decreased with increasing  $b_s$  in parenchyma. However, depending on the lesion type (hypo- versus hyperperfusion), a minimum of the parameter estimation error in the lesion was found for an intermediate  $b_s$  (300–400 s/mm<sup>2</sup> for cancer; 500 s/mm<sup>2</sup> for acute stroke, considering F only).

#### 4.1.1. Cancer

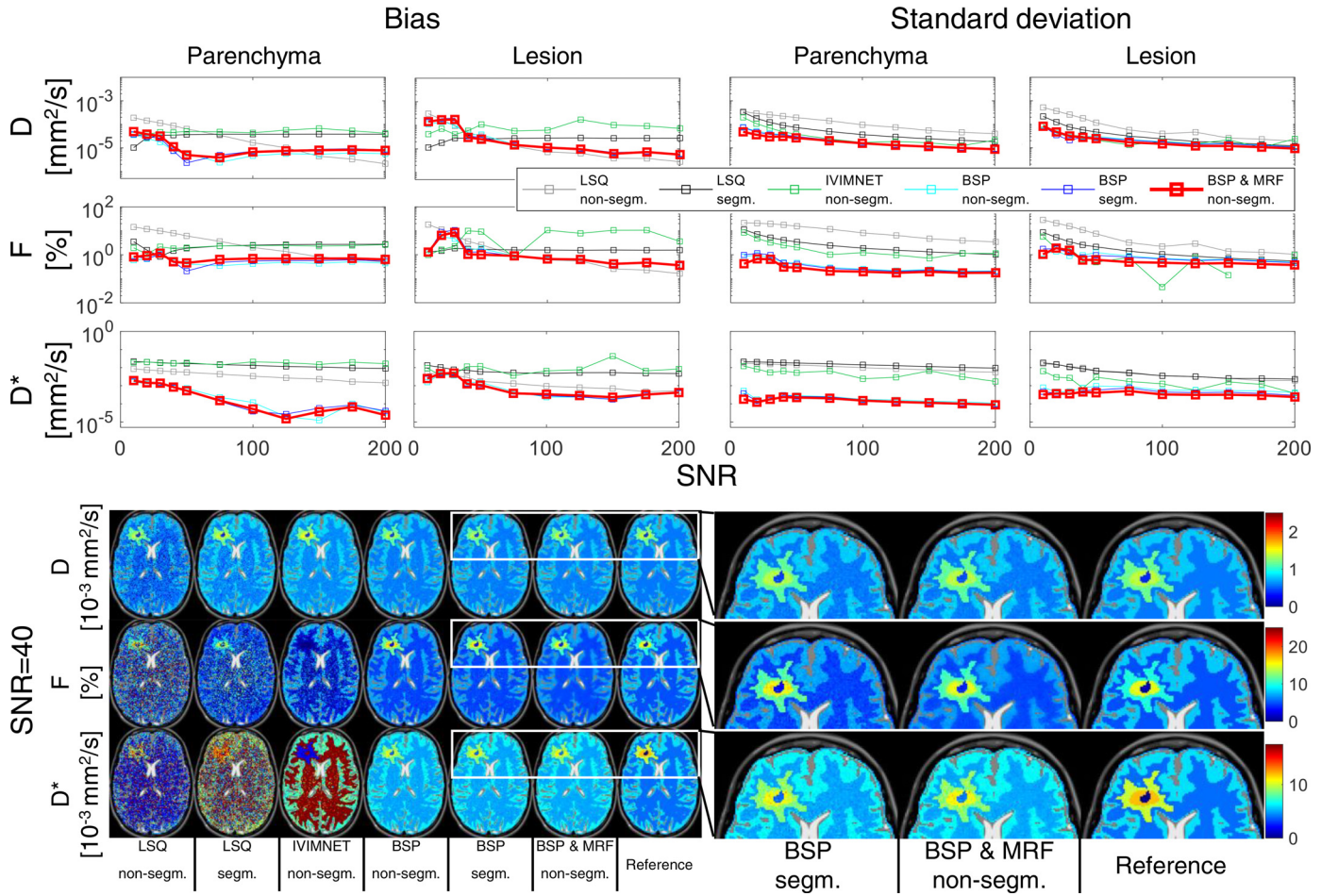
Fig. 1 shows results of the brain cancer scenario simulation. Both bias and standard deviation of the estimates in the lesion and parenchyma were among the lowest of all methods across all investigated SNRs. The errors (bias±standard deviation) of BSP & MRF at SNR=40 were ( $10^{-3}$  mm<sup>2</sup>/s / % /  $10^{-3}$  mm<sup>2</sup>/s):  $0.01\pm0.03/0.5\pm0.3/0.8\pm0.2$  and  $0.03\pm0.03/1.0\pm0.6/1.3\pm0.5$  in parenchyma and lesion respectively. The corresponding maps appeared similar to the reference maps, preserved the edges and allowed for clear delineation of the lesion in all IVIM parameter maps, while the LSQ methods partially obscured the lesion. The IVIMNET method exhibited contrast inversion of the perfusion parameters by underestimating F and D\* in the cancer lesion, whereby the latter was overestimated in the healthy parenchyma. The additional MRF prior effectively further removed remaining speckle-like noise in the central region.

The relative amount of voxels with  $\hat{R} > 1.1$  in the cancer simulation (SNR=40) were 0.0/0.0/0.0, 0.1/0.0/0.0 and 0.1/0.0/0.0% for BSP non-segm./BSP segm./BSP & MRF and D, F and D\*, respectively.

#### 4.1.2. Acute stroke

Fig. 2 shows results of the acute stroke scenario simulation. Both bias and standard deviation of the estimates in the lesion and parenchyma were among the lowest of all methods across all investigated SNRs. The errors (bias±standard deviation) of BSP & MRF at SNR=30 were ( $10^{-3}$  mm<sup>2</sup>/s / % /  $10^{-3}$  mm<sup>2</sup>/s):  $0.00\pm0.04/0.9\pm0.3/0.3\pm0.3$  and  $0.01\pm0.02/1.5\pm0.2/1.0\pm0.2$  in parenchyma and lesion, respectively. The corresponding maps appeared similar to the reference maps, preserved the edges and allowed for clear delineation of the lesion in all IVIM parameters, while the LSQ methods partially obscured the lesion. The IVIMNET method also allowed for clear delineation of the lesion in the F map, but exhibited contrast inversion for D\*: Parameter estimates in the lesion were close to the upper parameter constraint, while estimates in the parenchyma were smaller. The additional MRF prior effectively further removed remaining speckle-like noise in the central region. Smoothing was especially noticeable in the maps for F. The non-segmented BSP method exhibited a contrast inversion: perfusion fraction F was over-estimated, especially in the stroke lesion, where it was higher than in healthy parenchyma.

The relative amount of voxels with  $\hat{R} > 1.1$  in the acute stroke simulation (SNR=30) was 5.2/0.0/0.0, 35.4/0.0/0.0 and 9.0/0.0/0.0% for BSP non-segm./BSP segm./BSP & MRF and D, F and D\*, respectively.



**Fig. 1.** Simulation of cancer. Upper part: error (bias and standard deviation) of all investigated inference methods both in the cancer lesion and in healthy remaining parenchyma. Lower part: example parameter maps for all IVIM parameters and methods for a signal-to-noise ratio (SNR) of 40. The proposed method (BSP & MRF) yielded low bias, high precision and preserved edges in the parameter maps. LSQ=least squares, IVIMNET=deep neural network, BSP=Bayesian shrinkage prior, MRF=Markov random field, segm.=segmented.

#### 4.2. In-vivo data

An overview of median $\pm$ inter-quartile range of IVIM parameter estimates *in-vivo* for all investigated methods is presented in Table 2.

##### 4.2.1. Cancer

Mean SNR was  $43.4 \pm 16.4$  considering segmented voxels in cancer patients *in-vivo*.

An *in-vivo* example of cancer is shown in Fig. 3. A lesion on the right hemisphere could be seen with reduced central, but enhanced surrounding diffusivity (D). The maps of the LSQ methods appeared grainy and showed frequent outliers. The IVIMNET estimates yielded increased estimates for D\* compared to the other methods. The non-segmented BSP approach yielded relatively reduced D, increased F and reduced D\*. The proposed approach (BSP & MRF) allowed for delineation of the lesion and allowed to discriminate between central region (reduced diffusion, enhanced perfusion) and a surrounding region (enhanced diffusion, reduced perfusion).

CNR values of cancer patient data are reported in Fig. 4. The proposed method (BSP & MRF) yielded third-highest CNR (median $\pm$ inter-quartile range) for D after IVIMNET and segmented LSQ:  $5.9 \pm 6.2$ ,  $8.7 \pm 7.7$  and  $9.3 \pm 7.9$  respectively. CNR for F was highest using the proposed method ( $10.1 \pm 8.9$ ). All methods yielded

negative median CNR for D\*, where the proposed method produced a median CNR of  $-1.3 \pm 3.5$ .

Uncertainty values together with *in-vivo* example maps are shown in Fig. 5. Considering all methods, the proposed method (BSP & MRF) had lowest uncertainty. The median $\pm$ inter-quartile range was  $3.3 \pm 1.3\%$  for D and  $8.0 \pm 7.8\%$  for D\*, but highest for F: There, uncertainty was  $14.4 \pm 6.5\%$ . The inter-ventricular area had elevated uncertainty in D and corresponds to a region, where image registration had deformed the image more than in other areas. The uncertainty maps of perfusion estimates showed two spots of increased uncertainty. There, the perfusion was particularly low and perfusion parameter estimation hence difficult.

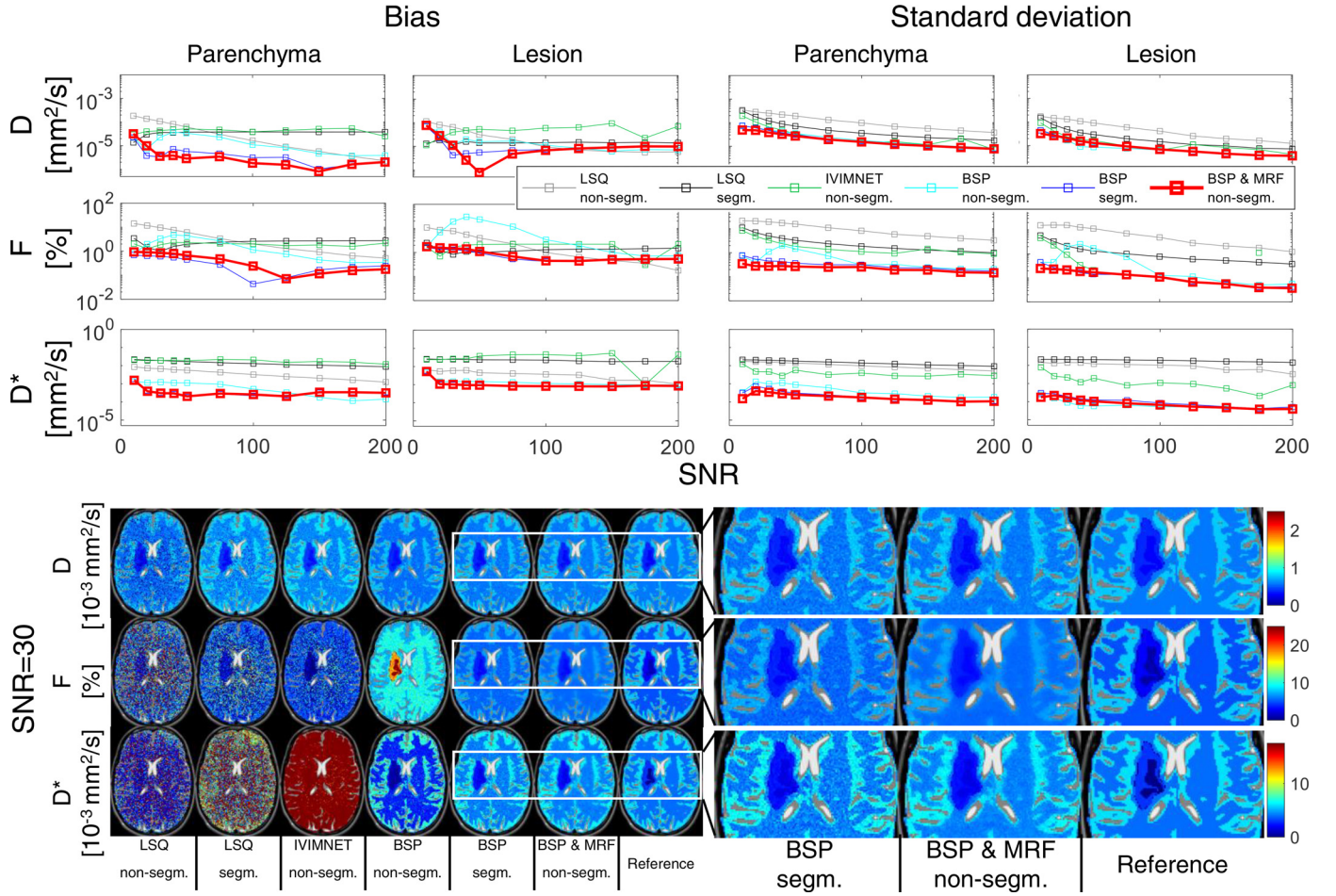
The relative amount of voxels with  $\hat{R} > 1.1$  in cancer *in-vivo* were  $6.1/0.0/0.0$ ,  $8.3/0.0/0.0$  and  $2.6/0.0/0.0\%$  for BSP non-segm./BSP segm./BSP & MRF and D, F and D\*, respectively.

##### 4.2.2. Acute stroke

Mean SNR was  $32.2 \pm 16.2$  considering segmented voxels in acute stroke patients *in-vivo*.

An *in-vivo* example of acute stroke is shown in Fig. 6. The parameter maps of D revealed a lesion with decreased diffusion in the central region of the right hemisphere. The maps derived from LSQ methods appeared grainy and showed frequent outliers. The IVIMNET estimates for D and F were comparable to the other methods. Also, the estimates for D\* were comparable to the other





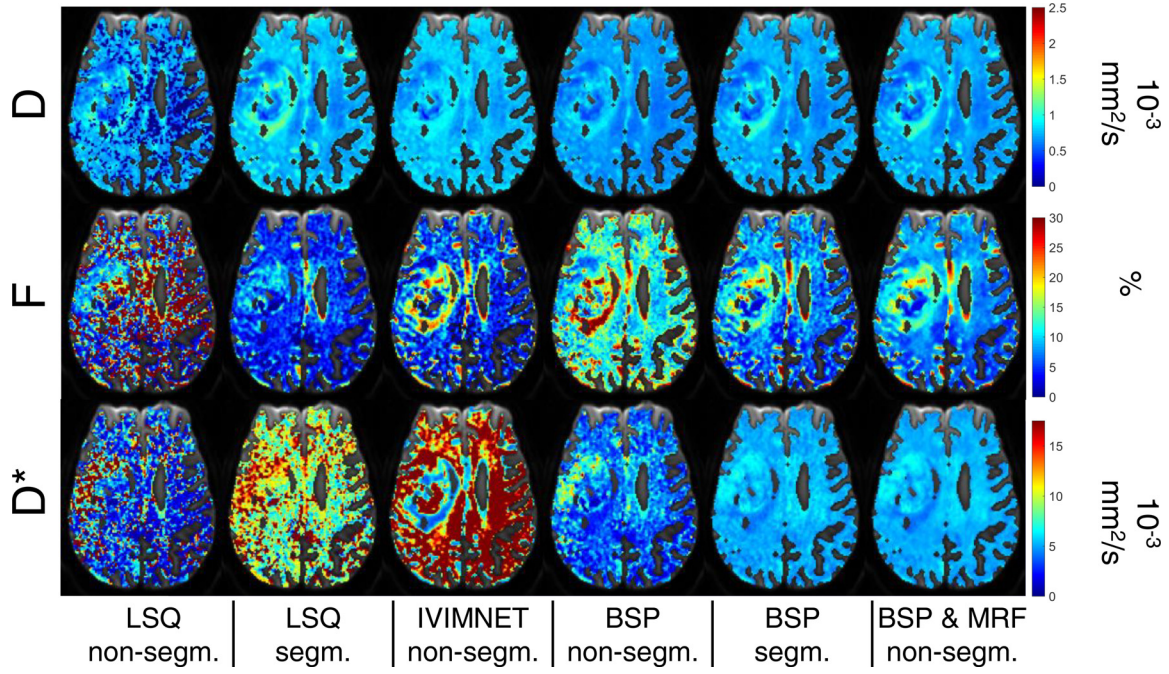
**Fig. 2.** Simulation of acute stroke. Upper part: error (bias and standard deviation) of all investigated inference methods both in the cancer lesion and in healthy remaining parenchyma. Lower part: example parameter maps for all IVIM parameters and methods for a signal-to-noise ratio (SNR) of 30. The proposed method (BSP & MRF) yielded low bias, high precision and preserved edges in the parameter maps. LSQ=least squares, IVIMNET=deep neural network, BSP=Bayesian shrinkage prior, MRF=Markov random field, segm.=segmented.

**Table 2**

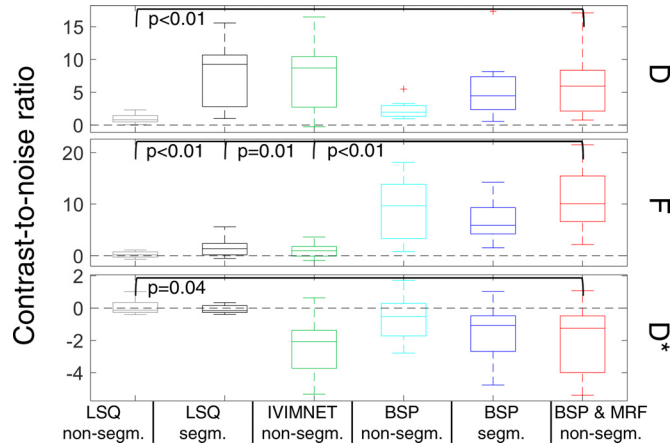
Median±inter-quartile range of *in-vivo* IVIM parameter estimates. The asterisk (\*) indicates that estimates in the lesion were significantly different from the ones in the control parenchyma ( $p < 0.05$ , Wilcoxon rank sum/Mann-Whitney  $U$  test with 5% significance level and Bonferroni correction for multiple testing). LSQ=least squares, IVIMNET=deep neural network, BSP=Bayesian shrinkage prior, MRF=Markov random field, segm.=segmented.

		Cancer			Acute stroke		
		All	Control	Lesion	All	Control	Lesion
D [ $10^{-3} \text{mm}^2/\text{s}$ ]	LSQ non-segm.	0.65±0.77	0.57±0.74	0.85±1.33*	0.62±0.75	0.53±0.71	0.33±0.42*
	LSQ segm.	0.78±0.15	0.78±0.13	1.77±0.73*	0.76±0.15	0.74±0.12	0.41±0.13*
	IVIMNET non-segm.	0.81±0.18	0.81±0.13	1.74±0.65*	0.78±0.16	0.77±0.13	0.47±0.12*
	BSP non-segm.	0.70±0.10	0.72±0.11	0.86±0.14*	0.67±0.12	0.64±0.10	0.40±0.11*
	BSP segm.	0.73±0.11	0.74±0.11	1.36±0.69*	0.71±0.12	0.69±0.10	0.40±0.11*
	BSP & MRF non-segm.	0.73±0.12	0.74±0.11	1.43±0.66*	0.70±0.13	0.68±0.10	0.41±0.12*
F [%]	LSQ non-segm.	17.0±59.6	21.6±65.3	43.0±77.7*	16.4±59.2	23.7±62.5	6.1±42.7*
	LSQ segm.	4.8±4.7	4.7±3.5	5.7±5.0*	5.0±5.4	5.2±4.0	2.4±3.4*
	IVIMNET non-segm.	5.4±6.7	5.5±5.4	6.7±4.3*	5.1±5.3	5.1±4.2	0.0±2.2*
	BSP non-segm.	11.34±10.32	9.29±7.42	72.70±43.27*	13.7±8.0	15.0±5.4	14.3±14.3
	BSP segm.	7.6±6.7	6.8±4.2	28.4±15.9*	8.0±5.5	8.3±3.7	3.4±2.4*
	BSP & MRF non-segm.	7.8±4.9	7.2±3.2	32.3±16.2*	8.6±4.0	8.8±2.1	3.8±1.6*
D* [ $10^{-3} \text{mm}^2/\text{s}$ ]	LSQ non-segm.	3.4±6.9	2.4±5.1	3.1±3.8*	3.5±9.5	2.4±5.2	2.9±11.0*
	LSQ segm.	9.6±7.6	8.7±6.0	9.2±3.6*	10.1±12.1	8.6±6.1	11.0±35.9*
	IVIMNET non-segm.	23.1±12.3	25.9±10.8	8.6±6.6*	24.5±14.1	22.3±13.4	40.0±25.5*
	BSP non-segm.	4.1±2.5	4.6±2.5	2.6±1.1*	3.4±3.4	3.2±1.8	1.2±1.7*
	BSP segm.	5.0±1.3	5.5±1.6	4.0±0.8*	4.8±2.4	4.5±1.4	3.6±1.8*
	BSP & MRF non-segm.	5.0±1.2	5.5±1.5	3.8±1.1*	4.8±2.2	4.6±1.0	3.7±1.5*





**Fig. 3.** IVIM parameters for an example cancer patient (63 years, female, glioblastoma with oligodendroglial component, grade 4) *in-vivo* with non-diffusion-weighted magnitude image in the background. A lesion with diffusion abnormalities could be identified in the right hemisphere with all fit methods. The proposed method (BSP & MRF) yielded visually relatively smooth parameter maps, where a clear visual delineation of the lesion was possible in all IVIM parameters. LSQ=least squares, IVIMNET=deep neural network, BSP=Bayesian shrinkage prior, MRF=Markov random field, segm.=segmented.



**Fig. 4.** Contrast-to-noise ratio (CNR) for all cancer *in-vivo* patient data. Positive values indicate that values in the lesion were higher than the ones in the control parenchyma. Significant differences among the proposed method (BSP & MRF) and the other methods were checked using a Kruskal-Wallis test and are indicated using a black bar. The proposed method (BSP & MRF) yielded a relatively high CNR for D and F compared to the other methods. LSQ=least squares, IVIMNET=deep neural network, BSP=Bayesian shrinkage prior, MRF=Markov random field, segm.=segmented.

methods in the central area, but higher than the largest value to display in peripheral areas. The non-segmented BSP approach yielded relatively reduced D, increased F and reduced D\*. The proposed approach (BSP & MRF) allowed for delineation of the lesion.

CNR values of acute stroke patient data are reported in Fig. 7. The proposed method (BSP & MRF) yielded a CNR (median $\pm$ inter-quartile range) for D of  $3.3\pm 2.1$ . CNR for F was highest using the proposed method ( $3.6\pm 1.8$ ). LSQ and IVIMNET methods yielded negative median CNR for D\*, where the proposed method produced a positive CNR of  $2.2\pm 4.8$ .

Uncertainty values together with *in-vivo* example maps are shown in Fig. 8. Considering all methods, the proposed method (BSP & MRF) had lowest uncertainty. The median $\pm$ inter-quartile range was  $3.8\pm 1.4\%$  for D and  $14.8\pm 4.5\%$  for F and  $19.6\pm 16.6\%$  for D\*. The uncertainty maps of perfusion estimates showed a spot of increased uncertainty in the stroke lesion. There, the perfusion was low due to the stroke and perfusion parameter estimation hence difficult.

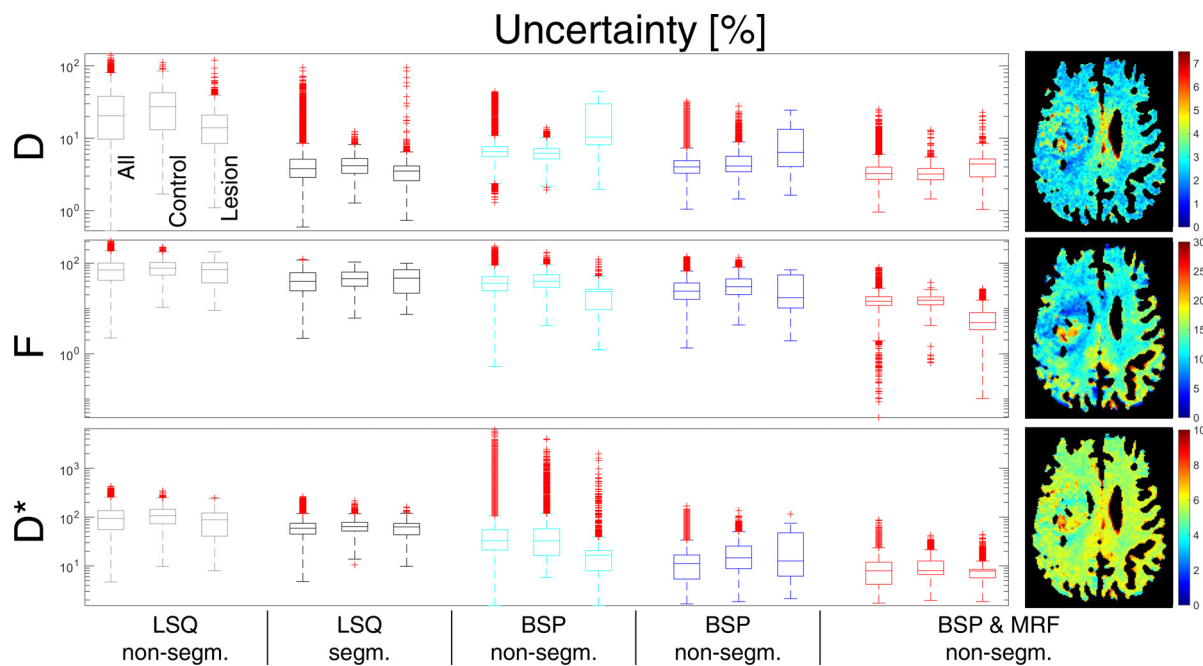
The relative amount of voxels with  $\hat{R} > 1.1$  in acute stroke *in-vivo* were 0.1/0.0/0.0, 0.1/0.0/0.0 and 0.0/0.0/0.0% for BSP non-segm./BSP segm./BSP & MRF and D, F and D\* respectively.

## 5. Discussion

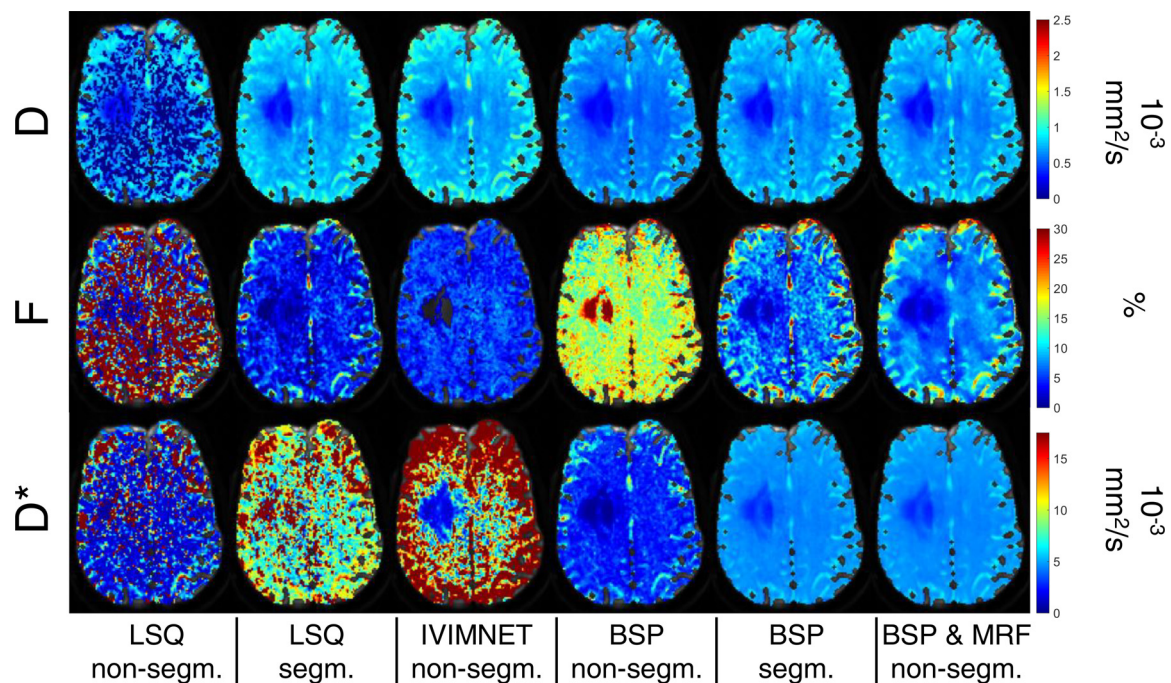
In this study, Bayesian inference using a combination of a hierarchical and a spatial prior was used, hyper-prior values were estimated from data using a model encompassing a segmented data likelihood, a hierarchical prior and Jeffreys hyper-prior. The proposed method has been shown to improve IVIM parameter estimation in cancer and acute stroke relative to conventional LSQ, IVIMNET and hierarchical Bayesian inference without spatial regularization or data likelihood segmentation.

### 5.1. IVIM estimates *in-vivo*

The presented results in terms of IVIM estimates *in-vivo* were within the range of estimates found in literature on brain cancer using segmented ( $b_0=200$  s/mm<sup>2</sup>) LSQ (Conklin et al., 2016; Federau et al., 2017; 2014a) and acute stroke (Conklin et al., 2016; Federau et al., 2014c; 2019). Interestingly, the perfusion fraction F using BSP & MRF was particularly high in cancer. Note that, on average, all methods apart from LSQ found a reduced D\* in the cancer lesion compared to the control region, which was also found previously (Federau et al., 2014a). These two findings suggest presence of numerous additional vessels but reduced blood flow velocity, indicating larger vessels in the cancer lesions. Supporting evidence



**Fig. 5.** Uncertainty as the coefficient of variation of the Markov chain Monte Carlo samples for all cancer patient data in-vivo. Boxplots contain all in-vivo cases, maps display one example (63 years, female, glioblastoma with oligodendroglial component, grade 4). LSQ uncertainty was estimated using flat priors, which accounted for the constraints used in the LSQ fits. The proposed method (BSP & MRF) had lowest uncertainty. LSQ=least squares, BSP=Bayesian shrinkage prior, MRF=Markov random field, segm.=segmented.



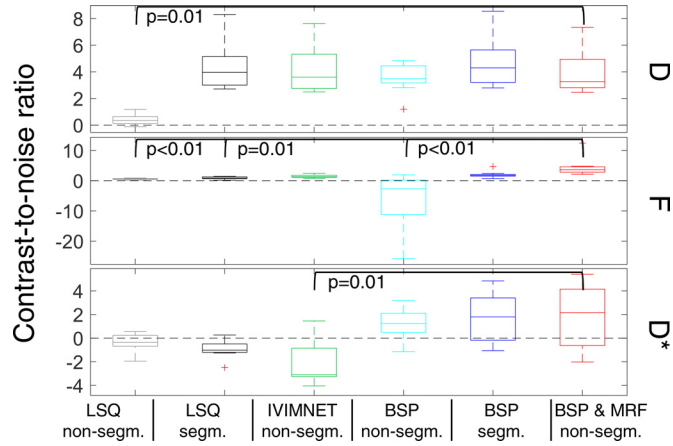
**Fig. 6.** IVIM parameters for an example patient with acute stroke (25 years, female, right middle cerebral artery region, imaging 5 days after symptom onset) in-vivo with non-diffusion-weighted magnitude image in the background. A lesion with reduced diffusion could be identified in the right hemisphere with most fit methods. The proposed method (BSP & MRF) yielded visually relatively smooth parameter maps, where a clear visual delineation of the lesion was possible in all IVIM parameters. LSQ=least squares, IVIMNET=deep neural network, BSP=Bayesian shrinkage prior, MRF=Markov random field, segm.=segmented.

was found in histological studies of glioblastoma tissue samples, where an increased amount of vessels was found (Pen et al., 2007).

In acute stroke, the perfusion fraction  $F$  was found to be about half of the estimates in the control region using the proposed method. This finding was also confirmed using conventional LSQ fits in this study and in more recent studies (Federau et al., 2019). The reduction in  $D^*$  was more subtle for

all methods apart from the LSQ and IVIMNET methods, where it paradoxically increased or remained virtually constant, respectively. This is physiologically implausible, because vessel occlusions lead to reduced blood flow in subsequent vasculature. Therefore, it is also perceivable that the ground truth values might be even smaller in the lesion, especially if perfusion is completely blocked.





**Fig. 7.** Contrast-to-noise ratio (CNR) for all acute stroke *in-vivo* patient data. Positive values indicate that values in the lesion are lower than ones in the control parenchyma. Significant differences among the proposed method (BSP & MRF) and the other methods were checked using a Kruskal-Wallis test and are indicated using a black bar. The proposed method (BSP & MRF) yielded a relatively high CNR compared to the other methods for all IVIM parameters. LSQ=least squares, IVIMNET=deep neural network, BSP=Bayesian shrinkage prior, MRF=Markov random field, segm.=segmented.

### 5.2. Least squares & deep neural networks

The LSQ methods exhibited large parameter errors and grainy-looking maps, especially the non-segmented version as frequently found in other studies (Pekar et al., 1992). The IVIMNET method yielded IVIM parameter estimates, which generally were comparable to the other methods for D and F, but D\* was frequently found to be larger than the ones from the other methods. Moreover, bias was found to be relatively high and not monotonically decreasing with SNR in simulations. Also, parameter contrast was inverted in cancer and acute stroke. It should be noted that results from deep learning methods depend on settings such as the cho-

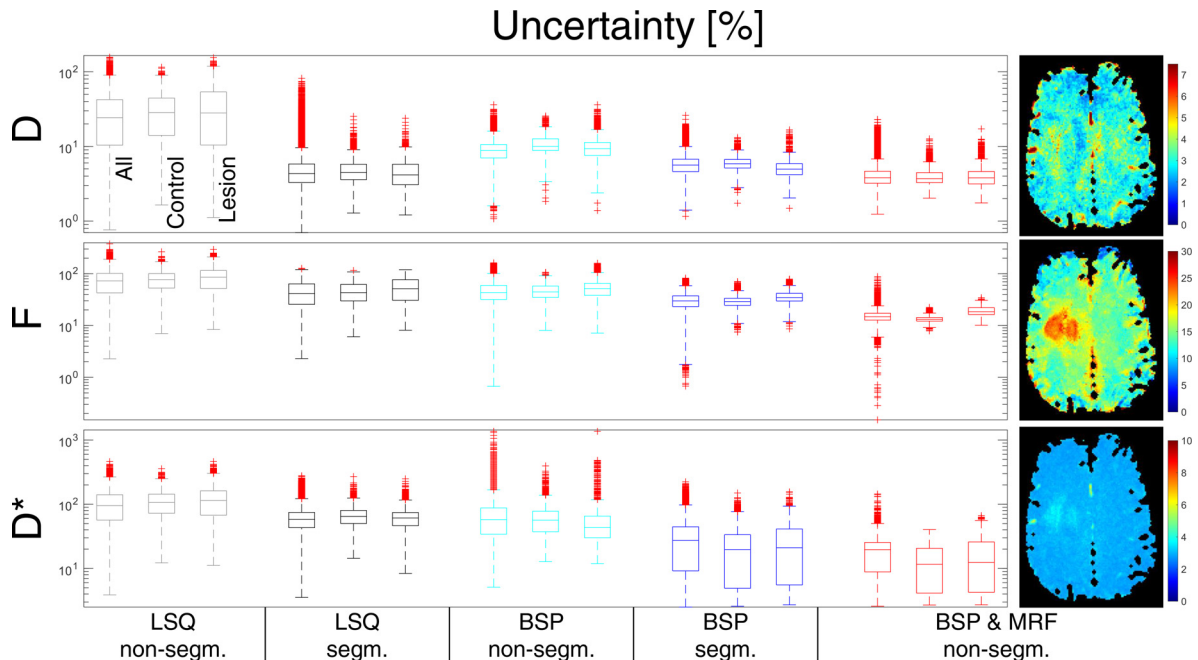
sen network architecture, activation functions and the numerous hyper-parameters, e.g. parameter constraints. Hence, it is perceivable that this method might benefit from modifications such as e.g. tighter parameter constraints, if prior knowledge is available for the considered application.

### 5.3. Data likelihood segmentation

The influence of data likelihood/model segmentation was noticeable both in LSQ and Bayesian methods. A higher cut-off than conventionally used ( $b_s = 200 \text{ s/mm}^2$ ) reduced bias for the proposed method (BSP & MRF) and also for the LSQ method (not shown here). In fact, it was chosen as high as possible given the acquisition  $b$ -values for the proposed method. This led to reduced error and more importantly to the prevention of false contrast inversion as occurring for acute stroke in the non-segmented BSP method (Orton et al., 2014; Spinner et al., 2017). Even though this non-segmented method did not introduce bias via data likelihood segmentation (due to residual perfusion influence on high  $b$ -values), the estimates were found to be more biased than the ones from the segmented BSP approach: F was found to be relatively large, while D\* was relatively small, both in simulations and *in-vivo*. The splitting  $b$ -value depends on the actual IVIM parameters to infer. Here, it was chosen to yield lowest error in healthy parenchyma, which was not always optimal for the lesion parameters. In principle, the choice of cut-off  $b$ -value could be made in a Bayesian approach, by treating the index of the splitting  $b$ -value as a random discrete variable modelled by a discrete distribution and imposing a Dirichlet prior. However, this would come at an additional computational cost.

### 5.4. Hierarchical prior

The hierarchical prior generally improved parameter estimation compared to standard LSQ fits. The partial pooling of estimates in the hierarchical model was hence an effective technique. However, that method also lead to estimates being “drawn” to the ROI mean



**Fig. 8.** Uncertainty as the coefficient of variation of the Markov chain Monte Carlo samples for acute stroke. LSQ uncertainty was estimated using flat priors, which accounted for the constraints used in the LSQ fits. The proposed method (BSP & MRF) had lowest uncertainty. LSQ=least squares, BSP=Bayesian shrinkage prior, MRF=Markov random field, segm.=segmented.



for all BSP methods, especially where perfusion was low. There, the influence of the data likelihood is small on the estimation of particularly  $F$  and  $D^*$ , while the prior is more influential. This resulted in increased perfusion estimates in low perfused areas. A similar underestimation of large ground-truth values was noted in simulations, however to a lesser extent. It should be noted however, that the resulting bias of BSP segm. and BSP & MRF was generally lower than that of non-regularized methods. The applied prior also modelled correlations among the parameters. It was found that *in-vivo* estimates using standard LSQ methods generally exhibited such correlation, i.e. both diffusion and perfusion varied jointly among the tissues. The hierarchical prior can potentially introduce bias in situations, where the correlations are strongly different across the ROI, e.g. small lesions with high diffusion, but low perfusion or vice versa. Such correlations can be removed from the prior by using three uni-variate Normal distributions without correlation instead. Likewise, the hyper-parameters can be set empirically, if sufficient prior knowledge is available, from e.g. large studies. Also, a more complex prior structure in the form of e.g. a Gaussian mixture model could be used to assign the voxels into several clusters, instead of only one. Different  $b_s$  could then be also used for the different clusters.

### 5.5. Jeffreys hyper-prior

The employed Jeffreys hyper-prior for the hierarchical prior is data-driven and thus frees the user from detailed prior knowledge about the IVIM parameters. This hyper-prior leads to a small determinant of the covariance matrix of the Normal prior and therefore to a narrow prior distribution. The resulting smoothing was particularly apparent if only one tissue type was investigated (Spinner et al., 2017). Here, the presence of different tissue types impeded this behavior to a certain extent. The prior was wide enough to accommodate the different tissue types, especially in the cancer scenario. However in acute stroke, the relatively few voxels in the lesion were falsely found to be of high perfusion fraction  $F$  for the BSP non-segm. method: the data likelihood influence becomes lower for smaller perfusion and the prior becomes dominant, which was not wide enough to accommodate these estimates. Also, this hyper-prior as used in the BSP non-segm. method often lead to bias in the perfusion parameters, despite not employing a segmented data likelihood. Actually, this problematic issue was overcome by using the segmented data likelihood as explained above in the segmented BSP method. Note that Jeffreys prior can be substituted with more informative empirical hyper-priors, which might be e.g. broader to reduce the pooling effect. Also, it is important to avoid the inclusion of many voxels with largely different IVIM parameters, such as CSF. This would lead to a broadening of the Normal prior, because of these outliers.

### 5.6. Spatial prior

The combination with a spatial MRF prior further reduced the estimation error, especially the residual speckle-like noise in the maps. Such a combined approach has recently been presented for dynamic contrast-enhanced cardiac MRI (Scannell et al., 2020), but not for IVIM to the best of our knowledge. The influence of this prior was particularly noticeable in the central brain region with reduced coil sensitivity and accordingly low SNR. The perfusion fraction  $F$  was particularly affected by this MRF prior. Its influence became more prominent if overall SNR was reduced. In general, the prior preserved edges and did not lead to over-smoothing, preserving small details *in-vivo*. The degree of smoothing can be controlled via the rate parameter  $\nu_{i,j}$ , which was set to 1 for all experiments in this study. Note also, that presence of many outliers, such as CSF, leads to an influence onto the neighboring voxels. Hence,

both the hierarchical and the MRF prior necessitate adequate segmentation with exclusion of CSF-rich areas. As an alternative,  $T_2$ -prepared acquisition sequences can be used, which suppress signal from CSF during read-out (Federau and O'Brien, 2015), a technique which was also found to increase perfusion contrast.

### 5.7. Limitations

Several limitations were present in this study. Simulations did not take into account the anisotropy of both diffusion (Le Bihan, 2003) and perfusion (Kubíková et al., 2018) in the brain. Spatial variations in the brain were simulated only on a tissue-type-level and not on a finer level. The lesion design was subjective and not based on a cohort, especially in terms of shape. Also, only one set of parameters and one lesion shape on one axial brain level was investigated, mainly due to computationally intensive MCMC sampling. Furthermore, only 8 instead of 32 surface coil channels were simulated, which might have influenced the SNR distribution and hence conditioning, especially in central brain regions. Likewise, *in-vivo* patients were heterogeneous and not stratified in terms of severity of pathology, lesion size and location. Comparisons with other perfusion measurement techniques such as arterial spin labeling or exogenous contrast agent-based methods were not performed, because of the lack of data thereof. It would be desirable to conduct and analyze further studies employing such reference methods in comparison with IVIM (Federau et al., 2019) and relate IVIM parameter estimates to clinical outcome (Federau et al., 2017). In a time-critical scenario, such as acute stroke diagnosis, fast execution is crucial. MCMC is computationally demanding and hence time-consuming, but several improvements are possible: optimization of the expensive MRF calculations, which are currently implemented as for-loops over the image dimensions and early stopping of the MCMC algorithm after e.g. a certain  $\hat{R}$  value is reached. The number of samples used for this study was more than sufficient for most datasets. In addition, MCMC can be executed in parallel on more processing units than shown in this study. Other methods have also been proposed for Bayesian inference from MR images, such as gradient-based variational inference or Langevin Monte Carlo (Dikaio, 2020), which might also allow for faster computation.

## 6. Conclusion

The proposed approach of combining a hierarchical with a spatial prior improved IVIM parameter estimation in both cancer and acute stroke compared to standard regression methods.

## Abbreviations

IVIM: intravoxel incoherent motion, TR: repetition time, TE: echo time, ROI: region-of-interest, LSQ: least squares, CNR: contrast-to-noise ratio, SNR: signal-to-noise ratio

## Declaration of Competing Interest

The 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.

## CRediT authorship contribution statement

**Georg Ralph Spinner:** Conceptualization, Formal analysis, Funding acquisition, Methodology, Project administration, Software, Validation, Visualization, Writing - original draft. **Christian Federau:** Data curation, Funding acquisition, Investigation, Resources, Supervision, Writing - review & editing. **Sebastian**

**Kozerke:** Funding acquisition, Resources, Supervision, Writing - review & editing.

## Acknowledgments

We would like to thank the developers of the IVIMNET method, Misha P.T. Kaandorp and Oliver J. Gurney-Champion for their advice in employing their method. Hannes Dillinger's help with the Linux workstation is appreciated. Furthermore, the reviewers' remarks, which improved this work are greatly appreciated.

## Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:[10.1016/j.media.2021.102144](https://doi.org/10.1016/j.media.2021.102144). Declarations and Appendix consisting of Supplemental Figures S1 & S2: Simulation procedure workflow & Parameter sweep of splitting  $b$ -value for the proposed method.

## References

- Gelman, A., Gilks, W.R., Roberts, G.O., 1997. Weak convergence and optimal scaling of random walk metropolis algorithms. *Ann. Appl. Probab.* 7 (1), 110–120. doi:[10.1214/aop/1034625254](https://doi.org/10.1214/aop/1034625254).
- Gelman, A., Rubin, D.B., 1992. Inference from iterative simulation using multiple sequences. *Stat. Sci.* 7 (4), 457–472. doi:[10.1214/ss/1177011136](https://doi.org/10.1214/ss/1177011136).
- 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 (5), 2175–2184. doi:[10.1002/mrm.25765](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 (1), 312–321. doi:[10.1002/mrm.27910](https://doi.org/10.1002/mrm.27910).
- Bardsley, J.M., 2012. Laplace-distributed increments, the laplace prior, and edge-preserving regularization. *J. Inverse Ill-Posed Probl.* 20 (3). doi:[10.1515/jip-2012-0017](https://doi.org/10.1515/jip-2012-0017).
- Berger, J.O., Bernardo, J.M., 1992. On the development of the reference prior method. *Bayesian Stat.* 4 (4), 35–60.
- Bisdas, S., Koh, T.S., Roder, C., Braun, C., Schittenhelm, J., Ernemann, U., Klose, U., 2013. Intravoxel incoherent motion diffusion-weighted MR imaging of gliomas: feasibility of the method and initial results. *Neuroradiology* 55 (10), 1189–1196. doi:[10.1007/s00234-013-1229-7](https://doi.org/10.1007/s00234-013-1229-7).
- Chan, T., Vese, L., 2001. Active contours without edges. *IEEE Trans. Image Process.* 10 (2), 266–277. doi:[10.1109/83.902291](https://doi.org/10.1109/83.902291).
- Conklin, J., Heyn, C., Roux, M., Cerny, M., Wintermark, M., Federau, C., 2016. A simplified model for intravoxel incoherent motion perfusion imaging of the brain. *Am. J. Neuroradiol.* 37 (12), 2251–2257. doi:[10.3174/ajnr.A4929](https://doi.org/10.3174/ajnr.A4929).
- , 1986. Bayesian inference and decision techniques: Essays in honor of Bruno de Finetti. In: De Finetti, B., Goel, P.K., Zellner, A. (Eds.), Number v. 6 In *Studies in Bayesian Econometrics and Statistics*. North-Holland; Sole distributors for the U.S.A. and Canada, Elsevier Science Pub. Co, Amsterdam ; New York : New York, N.Y., U.S.A.
- Dietrich, O., Raya, J.G., Reeder, S.B., Reiser, M.F., Schoenberg, S.O., 2007. Measurement of signal-to-noise ratios in MR images: influence of multichannel coils, parallel imaging, and reconstruction filters. *J. Magn. Reson. Imaging* 26 (2), 375–385. doi:[10.1002/jmri.20969](https://doi.org/10.1002/jmri.20969).
- Dikaio, N., 2020. Stochastic gradient langevin dynamics for joint parameterization of tracer kinetic models, input functions, and T1 relaxation-times from under-sampled k-space DCE-MRI. *Med. Image Anal.* 62, 101690. doi:[10.1016/j.media.2020.101690](https://doi.org/10.1016/j.media.2020.101690).
- Dyvorner, H.A., Galea, N., Nevers, T., Fiel, M.I., Carpenter, D., Wong, E., Orton, M., de Oliveira, A., Feiweier, T., Vachon, M.-L., Babb, J.S., Taouli, B., 2013. Diffusion-weighted imaging of the liver with multiple b values: effect of diffusion gradient polarity and breathing acquisition on image quality and intravoxel incoherent motion parameters—a pilot study. *Radiology* 266 (3), 920–929. doi:[10.1148/radiol.12120686](https://doi.org/10.1148/radiol.12120686).
- Federau, C., 2017. Intravoxel incoherent motion MRI as a means to measure *in vivo* perfusion: a review of the evidence. *NMR Biomed.* 30 (11), e3780. doi:[10.1002/nbm.3780](https://doi.org/10.1002/nbm.3780).
- Federau, C., Cerny, M., Roux, M., Mosimann, P.J., Maeder, P., Meuli, R., Wintermark, M., 2017. IVIM Perfusion fraction is prognostic for survival in brain glioma. *Clin. Neuroradiol.* 27 (4), 485–492. doi:[10.1007/s00062-016-0510-7](https://doi.org/10.1007/s00062-016-0510-7).
- Federau, C., Maeder, P., O'Brien, K., Browaeys, P., Meuli, R., Hagmann, P., 2012. Quantitative measurement of brain perfusion with intravoxel incoherent motion MR imaging. *Radiol.* 265 (3), 874–881. doi:[10.1148/radiol.12120584](https://doi.org/10.1148/radiol.12120584).
- Federau, C., Meuli, R., O'Brien, K., Maeder, P., Hagmann, P., 2014a. Perfusion measurement in brain gliomas with intravoxel incoherent motion MRI. *Am. J. Neuroradiol.* 35 (2), 256–262. doi:[10.3174/ajnr.A3686](https://doi.org/10.3174/ajnr.A3686).
- Federau, C., O'Brien, K., 2015. Increased brain perfusion contrast with T2-prepared intravoxel incoherent motion (T2prep IVIM) MRI. *NMR Biomed.* 28 (1), 9–16. doi:[10.1002/nbm.3223](https://doi.org/10.1002/nbm.3223).
- Federau, C., O'Brien, K., Meuli, R., Hagmann, P., Maeder, P., 2014b. Measuring brain perfusion with intravoxel incoherent motion (IVIM): initial clinical experience. *J. Magn. Reson. Imaging* 39 (3), 624–632. doi:[10.1002/jmri.24195](https://doi.org/10.1002/jmri.24195).
- Federau, C., Sumer, S., Becce, F., Maeder, P., O'Brien, K., Meuli, R., Wintermark, M., 2014c. Intravoxel incoherent motion perfusion imaging in acute stroke: initial clinical experience. *Neuroradiology* 56 (8), 629–635. doi:[10.1007/s00234-014-1370-y](https://doi.org/10.1007/s00234-014-1370-y).
- Federau, C., Wintermark, M., Christensen, S., Mlynash, M., Marcellus, D.G., Zhu, G., Martin, B.W., Lansberg, M.G., Albers, G.W., Heit, J.J., 2019. Collateral blood flow measurement with intravoxel incoherent motion perfusion imaging in hyperacute brain stroke. *Neurology* 92 (21), e2462. doi:[10.1212/WNL.00000000000007538](https://doi.org/10.1212/WNL.00000000000007538).
- Freiman, M., Perez-Rossello, J.M., Callahan, M.J., Voss, S.D., Ecklund, K., Mulkern, R.V., Warfield, S.K., 2013. Reliable estimation of incoherent motion parametric maps from diffusion-weighted MRI using fusion bootstrap moves. *Med. Image Anal.* 17 (3), 325–336. doi:[10.1016/j.media.2012.12.001](https://doi.org/10.1016/j.media.2012.12.001).
- Geman, S., Geman, D., 1984. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Trans. Pattern Anal. Mach. Intell. PAMI-6* (6), 721–741. doi:[10.1109/TPAMI.1984.4767596](https://doi.org/10.1109/TPAMI.1984.4767596).
- Gudbjartsson, H., Patz, S., 1995. The Rician distribution of noisy MRI data. *Magn. Reson. Med.* 34 (6), 910–914. doi:[10.1002/mrm.1910340618](https://doi.org/10.1002/mrm.1910340618).
- Guerquin-Kern, M., Lejeune, L., Pruessmann, K.P., Unser, M., 2012. Realistic analytical phantoms for parallel magnetic resonance imaging. *IEEE Trans. Med. Imaging* 31 (3), 626–636. doi:[10.1109/TMI.2011.2174158](https://doi.org/10.1109/TMI.2011.2174158).
- Gurney-Champion, O.J., Klaassen, R., Froeling, M., Barbieri, S., Stoker, J., Engelbrecht, M.R.W., Wilmink, J.W., Besseling, 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 (4), e0194590. doi:[10.1371/journal.pone.0194590](https://doi.org/10.1371/journal.pone.0194590).
- Gustafsson, O., Montelius, M., Starck, G., Ljungberg, M., 2018. Impact of prior distributions and central tendency measures on Bayesian intravoxel incoherent motion model fitting. *Magn. Reson. Med.* 79 (3), 1674–1683. doi:[10.1002/mrm.26783](https://doi.org/10.1002/mrm.26783).
- Jeffreys, H., 1946. An invariant form for the prior probability in estimation problems. *Proc. R. Soc. Lond. Ser. A Math. Phys. Sci.* 186 (1007), 453–461. doi:[10.1098/rspa.1946.0056](https://doi.org/10.1098/rspa.1946.0056).
- 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.* ahead of print (n/a). doi:[10.1002/mrm.28852](https://doi.org/10.1002/mrm.28852).
- Kass, R.E., Carlin, B.P., Gelman, A., Neal, R.M., 1998. Markov chain monte carlo in practice: a roundtable discussion. *Am. Stat.* 52 (2), 93–100. doi:[10.1080/00031305.1998.10480547](https://doi.org/10.1080/00031305.1998.10480547).
- Keil, V.C., Mädler, B., Gielen, G.H., Pintea, B., Hiththetiya, K., Gaspranova, A.R., Gieseke, J., Simon, M., Schild, H.H., Hadzadeh, D.R., 2017. Intravoxel incoherent motion MRI in the brain: impact of the fitting model on perfusion fraction and lesion differentiability. *J. Magn. Reson. Imaging* 46 (4), 1187–1199. doi:[10.1002/jmri.25615](https://doi.org/10.1002/jmri.25615).
- Kubíková, T., Kochová, P., Tomášek, P., Witter, K., Tonar, Z., 2018. Numerical and length densities of microvessels in the human brain: correlation with preferential orientation of microvessels in the cerebral cortex, subcortical grey matter and white matter, pons and cerebellum. *J. Chem. Neuroanat.* 88, 22–32. doi:[10.1016/j.jchemneu.2017.11.005](https://doi.org/10.1016/j.jchemneu.2017.11.005).
- Le Bihan, D., 2003. Looking into the functional architecture of the brain with diffusion MRI. *Nat. Rev. Neurosci.* 4 (6), 469–480. doi:[10.1038/nrn1119](https://doi.org/10.1038/nrn1119).
- Le Bihan, D., 2019. What can we see with IVIM MRI? *Neuroimage* 187, 56–67. doi:[10.1016/j.neuroimage.2017.12.062](https://doi.org/10.1016/j.neuroimage.2017.12.062).
- Le Bihan, D., Breton, E., Lallemand, D., Grenier, P., Cabanis, E., Laval-Jeantet, M., 1986. MR Imaging of intravoxel incoherent motions: application to diffusion and perfusion in neurologic disorders. *Radiology* 161 (2), 401–407. doi:[10.1148/radiology.161.2.3763909](https://doi.org/10.1148/radiology.161.2.3763909).
- Lemke, A., Laun, F.B., Klau, M., Re, T.J., Simon, D., Delorme, S., Schad, L.R., Stieltjes, B., 2009. Differentiation of pancreas carcinoma from healthy pancreatic tissue using multiple b-values: comparison of apparent diffusion coefficient and intravoxel incoherent motion derived parameters. *Invest. Radiol.* 44 (12), 769–775. doi:[10.1097/RLI.0b013e3181b62271](https://doi.org/10.1097/RLI.0b013e3181b62271).
- Lemke, A., Stieltjes, B., Schad, L.R., Laun, F.B., 2011. Toward an optimal distribution of b values for intravoxel incoherent motion imaging. *Magn. Reson. Imaging* 29 (6), 766–776. doi:[10.1016/j.mri.2011.03.004](https://doi.org/10.1016/j.mri.2011.03.004).
- Luciani, A., Vignaud, A., Cavet, M., Tran Van Nhieu, J., Mallat, A., Ruel, L., Laurent, A., Deux, J.-F., Brugieres, P., Rahmouni, A., 2008. Liver cirrhosis: intravoxel incoherent motion MR imaging—pilot study. *Radiology* 249 (3), 891–899. doi:[10.1148/radiol.2493080080](https://doi.org/10.1148/radiol.2493080080).
- Meeus, E.M., Novak, J., Withey, S.B., Zarinabad, N., Dehghani, H., Peet, A.C., 2017. Evaluation of intravoxel incoherent motion fitting methods in low-perfused tissue. *J. Magn. Reson. Imaging* 45 (5), 1325–1334. doi:[10.1002/jmri.25411](https://doi.org/10.1002/jmri.25411).
- Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H., Teller, E., 1953. Equation of state calculations by fast computing machines. *J. Chem. Phys.* 21 (6), 1087–1092. doi:[10.1063/1.1699114](https://doi.org/10.1063/1.1699114).
- Moulin, K., Croisille, P., Feiweier, T., Delattre, B.M.A., Wei, H., Robert, B., Beuf, O., Viallon, M., 2016. In vivo free-breathing DTI and IVIM of the whole human heart using a real-time slice-followed SE-EPI navigator-based sequence: a reproducibility study in healthy volunteers. *Magn. Reson. Med.* 76 (1), 70–82. doi:[10.1002/mrm.25852](https://doi.org/10.1002/mrm.25852).

- 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 (5), 642–647. doi:[10.1002/mrm.1910290510](https://doi.org/10.1002/mrm.1910290510).
- Notohamiprodjo, M., Chandarana, H., Mikheev, A., Rusinek, H., Grinstead, J., Feiweier, T., Raya, J.G., Lee, V.S., Sigmund, E.E., 2015. Combined intravoxel incoherent motion and diffusion tensor imaging of renal diffusion and flow anisotropy. *Magn. Reson. Med.* 73 (4), 1526–1532. doi:[10.1002/mrm.25245](https://doi.org/10.1002/mrm.25245).
- Orton, M.R., Collins, D.J., Koh, D.-M., Leach, M.O., 2014. Improved intravoxel incoherent motion analysis of diffusion weighted imaging by data driven Bayesian modeling. *Magn. Reson. Med.* 71 (1), 411–420. doi:[10.1002/mrm.24649](https://doi.org/10.1002/mrm.24649).
- Otsu, N., 1979. A threshold selection method from gray-level histograms. *IEEE Trans. Syst. Man Cybern.* 9 (1), 62–66. doi:[10.1109/TSMC.1979.4310076](https://doi.org/10.1109/TSMC.1979.4310076).
- Pekar, J., Moonen, C.T.W., van Zijl, P.C.M., 1992. On the precision of diffusion/perfusion imaging by gradient sensitization. *Magn. Reson. Med.* 23 (1), 122–129. doi:[10.1002/mrm.1910230113](https://doi.org/10.1002/mrm.1910230113).
- Pen, A., Moreno, M.J., Martin, J., Stanimirovic, D.B., 2007. Molecular markers of extracellular matrix remodeling in glioblastoma vessels: microarray study of laser-captured glioblastoma vessels. *Glia* 55 (6), 559–572. doi:[10.1002/glia.20481](https://doi.org/10.1002/glia.20481).
- Puig, J., Sánchez-González, J., Blasco, G., Daunis-i-Estadella, P., Federau, C., Alberich-Bayarri, Á., Biarnes, C., Nael, K., Essig, M., Jain, R., Wintermark, M., Pedraza, S., 2016. Intravoxel incoherent motion metrics as potential biomarkers for survival in glioblastoma. *PLoS ONE* 11 (7), e0158887. doi:[10.1371/journal.pone.0158887](https://doi.org/10.1371/journal.pone.0158887).
- Reeder, S.B., Wintersperger, B.J., Dietrich, O., Lanz, T., Greiser, A., Reiser, M.F., Glazer, G.M., Schoenberg, S.O., 2005. Practical approaches to the evaluation of signal-to-noise ratio performance with parallel imaging: application with cardiac imaging and a 32-channel cardiac coil. *Magn. Reson. Med.* 54 (3), 748–754. doi:[10.1002/mrm.20636](https://doi.org/10.1002/mrm.20636).
- Roemer, P.B., Edelstein, W.A., Hayes, C.E., Souza, S.P., Mueller, O.M., 1990. The NMR phased array. *Magn. Reson. Med.* 16 (2), 192–225. doi:[10.1002/mrm.1910160203](https://doi.org/10.1002/mrm.1910160203).
- Rydhög, A.S., van Osch, M.J.P., Lindgren, E., Nilsson, M., Lätt, J., Ståhlberg, F., Wirestam, R., Knutsson, L., 2014. Intravoxel incoherent motion (IVIM) imaging at different magnetic field strengths: what is feasible? *Magn. Reson. Imaging* 32 (10), 1247–1258. doi:[10.1016/j.mri.2014.07.013](https://doi.org/10.1016/j.mri.2014.07.013).
- Scannell, C.M., Chiribiri, A., Villa, A.D., Breeuwer, M., Lee, J., 2020. Hierarchical Bayesian myocardial perfusion quantification. *Med. Image Anal.* 60, 101611. doi:[10.1016/j.media.2019.101611](https://doi.org/10.1016/j.media.2019.101611).
- Spinner, G.R., Schmidt, J.F.M., von Deuster, C., Federau, C., Stoeck, C.T., Kozerke, S., 2018. Enhancing intravoxel incoherent motion parameter mapping in the brain using k-b PCA. *NMR Biomed.* 31 (12), e4008. doi:[10.1002/nbm.4008](https://doi.org/10.1002/nbm.4008).
- Spinner, G.R., von Deuster, C., Tezcan, K.C., Stoeck, C.T., Kozerke, S., 2017. Bayesian intravoxel incoherent motion parameter mapping in the human heart. *J. Cardiovasc. Magn. Reson.* 19 (1), 85. doi:[10.1186/s12968-017-0391-1](https://doi.org/10.1186/s12968-017-0391-1).
- Stejskal, E.O., Tanner, J.E., 1965. Spin diffusion measurements: spin echoes in the presence of a time-dependent field gradient. *J. Chem. Phys.* 42 (1), 288–292. doi:[10.1063/1.1695690](https://doi.org/10.1063/1.1695690).
- Suo, S., Cao, M., Zhu, W., Li, L., Li, J., Shen, F., Zu, J., Zhou, Z., Zhuang, Z., Qu, J., Chen, Z., Xu, J., 2016. Stroke assessment with intravoxel incoherent motion diffusion-weighted MRI. *NMR Biomed.* 29 (3), 320–328. doi:[10.1002/nbm.3467](https://doi.org/10.1002/nbm.3467).
- Togao, O., Hiwatashi, A., Yamashita, K., Kikuchi, K., Mizoguchi, M., Yoshimoto, K., Suzuki, S.O., Iwaki, T., Obara, M., Van Cauteren, M., Honda, H., 2016. Differentiation of high-grade and low-grade diffuse gliomas by intravoxel incoherent motion MR imaging. *Neurooncology* 18 (1), 132–141. doi:[10.1093/neuonc/nov147](https://doi.org/10.1093/neuonc/nov147).
- Vidić, I., Jerome, N.P., Bathen, T.F., Goa, P.E., While, P.T., 2019. Accuracy of breast cancer lesion classification using intravoxel incoherent motion diffusion-weighted imaging is improved by the inclusion of global or local prior knowledge with Bayesian methods. *J. Magn. Reson. Imaging* 50 (5), 1478–1488. doi:[10.1002/jmri.26772](https://doi.org/10.1002/jmri.26772).
- Vishnevskiy, V., Gass, T., Szekely, G., Tanner, C., Goksel, O., 2017. Isotropic total variation regularization of displacements in parametric image registration. *IEEE Trans. Med. Imaging* 36 (2), 385–395. doi:[10.1109/TMI.2016.2610583](https://doi.org/10.1109/TMI.2016.2610583).
- Wetscherek, A., Stieltjes, B., Laun, F.B., 2015. Flow-compensated intravoxel incoherent motion diffusion imaging. *Magn. Reson. Med.* 74 (2), 410–419. doi:[10.1002/mrm.25410](https://doi.org/10.1002/mrm.25410).
- 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 (6), 2373–2387. doi:[10.1002/mrm.26598](https://doi.org/10.1002/mrm.26598).
- Wu, W.-C., Chen, Y.-F., Tseng, H.-M., Yang, S.-C., My, P.-C., 2015. Caveat of measuring perfusion indexes using intravoxel incoherent motion magnetic resonance imaging in the human brain. *Eur. Radiol.* 25 (8), 2485–2492. doi:[10.1007/s00330-015-3655-x](https://doi.org/10.1007/s00330-015-3655-x).
- Yamada, I., Aung, W., Himeno, Y., Nakagawa, T., Shibuya, H., 1999. Diffusion coefficients in abdominal organs and hepatic lesions: evaluation with intravoxel incoherent motion echo-planar MR imaging. *Radiology* 210 (3), 617–623. doi:[10.1148/radiology.210.3.r99fe17617](https://doi.org/10.1148/radiology.210.3.r99fe17617).