MRI of pancreatic cancer for radiotherapy

Gurney-Champion, O.J.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
CHAPTER 8

Best fits for DWI

How best to fit diffusion-weighted magnetic resonance imaging data of pancreatic cancer patients

Oliver J Gurney-Champion
Remy Klaassen
Martijn Froeling
Sebastiano Barbieri
Jaap Stoker
Marc RW Engelbrecht
Johanna W Wilmink
Marc G Besselink
Arjan Bel
Hanneke WM van Laarhoven
Aart J Nederveen

Submitted
Abstract

Purpose: To compare the performance of six intravoxel incoherent motion (IVIM) model fitting algorithms and seven diffusion-weighted imaging (DWI) models for pancreatic cancer imaging.

Methods: The local medical ethics committee approved this prospective study. Patients gave written informed consent. Between October 2014 and March 2016, diffusion-weighted imaging (DWI) data were acquired in 14 pancreatic cancer patients (mean age: 68 years old) during two visits. For six IVIM fit algorithms (IVIM-free, IVIM-adaptive, IVIM-Bayesian-log, IVIM-Bayesian-lin, IVIM-MLE and IVIM-fixed) and seven DWI models (IVIM, mono-exp, mono-exp-2, tri-exp, stretched-exp, Gaussian and kurtosis) the (a) goodness of fit (represented by adjusted $R^2$); (b) uniqueness of fit parameters (Spearman’s rho); (c) precision (within-subject coefficient of variation, wCV); and (d) contrast between tumor and non-tumorous pancreatic tissue were assessed. A Wilcoxon signed-rank test ($\alpha = 0.05$) was performed between the best scoring parameter and remaining parameters (a, c, d), both for fit algorithms and DWI models.

Results: IVIM-Bayesian-lin and IVIM-fixed were overall the best IVIM fit algorithms. From the DWI models, Gaussian, kurtosis and monoexponential scored poorest. Excluding those, DWI models had high goodness of fit (a: $R^2 > 0.93$) and unique parameters (b). The perfusion fraction $f_2$ (tri-exp) had significantly higher tumor contrast (d) than all other parameters. Only the stretched-exp model had two precise (c: wCV < 15%) parameters, of which its perfusion-related parameter was significantly more precise than all other perfusion-related parameters.

Conclusion: IVIM-Bayesian-lin and IVIM-fixed are the best fit algorithms. The IVIM model scored relatively high in all tests, but for specific purposes, other models may be preferred.
Introduction

The intravoxel incoherent motion (IVIM) model for diffusion-weighted imaging (DWI) data obtained by magnetic resonance imaging (MRI) bears much promise as a tool to visualize tumors and monitor treatment response (e.g., in radiotherapy or chemotherapy) [1-3]. Contrary to the classical DWI model, in which signal attenuation is modeled monoexponentially as a function of diffusion-weighting (b-value), the IVIM model predicts a biexponential decay, probing both tissue diffusion and perfusion. Since the introduction of the IVIM model [4], the non-monoexponential behavior of DWI data in the pancreas was confirmed in multiple studies [1-3] and related to the interplay between diffusion and perfusion [5, 6]. Consequently, the IVIM model has been used to distinguish pancreatic cancer and surrounding tissue [3], characterized lesions in the pancreas [1, 2] and enabled treatment response monitoring in various other organs [7, 8].

Due to the limited precision of the IVIM parameters [9], multiple algorithms have been used to fit IVIM data. It was shown in simulations and healthy volunteer data that a Bayesian fit gives the most accurate and precise results in abdominal tissue [10]. However, thus far no analyses were done in pancreatic cancer patients. In these patients, fitting the data may be more challenging due to the limited size of the tumor compared to the entire organ, less perfusion [1-3], and echo planar imaging (EPI) artifacts that occur as a result of e.g. intratumoral fiducials [11] or biliary stents [12] close to the tumor.

The IVIM model is an approximation of the signal decay [13], and some of the assumptions underlying the IVIM model are not met in the pancreas [14]. Incoherent dephasing spins that result from different blood flow velocities within larger vessels can necessitate a third exponent in the IVIM model. Furthermore, rather than discrete (pseudo-) diffusion speeds, as used in the IVIM model, a smooth distribution of diffusion coefficients, reflecting the variation in cell sizes and vessel lengths within the voxel may be modeled. This was done in the stretched exponential [15] and Gaussian models [16].

Currently, there is no consensus on which fit algorithm for IVIM modeling and which DWI model should be used for different applications. We hypothesize that less common algorithms and fit models may outperform the most commonly used least squares fit of the mono-exponential and IVIM models for specific purposes. Hence, the objective of this exploratory study was to compare the performance of six IVIM model fitting algorithms and seven DWI models for pancreatic cancer imaging.
CHAPTER 8

Materials and Methods

This prospective study (NCT01995240) was approved by the local ethics committee, and all patients gave written informed consent. Inclusion criteria were: locally advanced or metastatic pancreatic ductal adenocarcinoma, normal kidney function (eGFR > 60 mL/min/1.73m²) and no contraindication to undergoing MRI scanning. Sixteen consecutive patients fulfilling these criteria and willing to participate were included. Patients were scanned on a 3 T scanner (Philips Ingenia, Best, The Netherlands) between October 2014 and March 2016 at the Academic Medical Center in Amsterdam. One patient dropped out between scan sessions, and for one patient, the scans were stopped due to patient discomfort. Thus, data from fourteen patients were analyzed (eight females, mean age 67 years old, range 52–78, six males, mean age 70 years old, range 56–77). Data from nine of these patients were published previously [17].

Data acquisition

All patients were scanned three times during two separate sessions (average: 4.5 days apart, range: 1–8 days). To minimize bowel motion hyoscine bromide (Buscopan, Boehringer, Ingelheim, Germany; 20 mg IV) was administered directly before the first DWI acquisition in each session. The data from the acquisition without administration of hyoscine bromide were used for the intra-session analysis only.

For each patient, we acquired 2D multi-slice diffusion-weighted EPI data and contrast-enhanced (CE) T1-weighted multi-echo spoiled gradient echo (T1W GE) data with Dixon reconstruction (see Table 8.1 for detailed imaging parameters). The T1W GE data were acquired 35 seconds after Gadovist 1.0 (Bayer Healthcare, Leverkusen, Germany) administration (0.1ml/kg; 5ml/s, followed by 15 ml saline flush). Per b-value, data were acquired in isotropic distributed directions.

Post processing

All data analyses, fitting and statistical tests were performed in Matlab 2013a (MathWorks, Natick, U.S.A.), except for the IVIM-Bayesian-log fit, which was implemented in DTITools for Mathematica [18], Mathematica 10.4.1 (Wolfram Research, Champaign, U.S.A.).

All DWI images were denoised using a Rician adaptive non-local means filter [19] and registered in Elastix [20, 21] (Appendix I for details).

First, we tested the most common fitting algorithms for the IVIM model: IVIM-free, IVIM-adaptive [22], IVIM-Bayesian-log [23, 24], IVIM-Bayesian-lin [10], IVIM-MLE
[25] and IVIM-fixed (Appendix II for details). All IVIM model fit algorithms converted the IVIM signal fractions into volume fractions [6], using a TE = 45 ms and assuming a TR = 5000 ms (typical respiratory cycle), T1 = 725 ms and T2 = 43 ms for the pancreas and T1 = 1932 ms and T2 = 275 ms for blood [26, 27].

Then, we tested six additional common DWI models using the Levenberg-Marquardt least squares fit: the mono-exp, mono-exp-2, tri-exp, stretched-exp, Gaussian and kurtosis model (Appendix II for details). These were also compared to the two Levenberg-Marquardt least squares IVIM fits: IVIM-free and IVIM-fixed.

An abdominal radiologist (M.R.W.E., 9 years' experience) and an abdominal imaging researcher (R.K. 3 years' experience) drew regions of interest (ROIs) in consensus using 3D Slicer [28]. For each patient, two ROIs were created, one containing pancreatic tumor tissue and one containing non-tumorous pancreatic tissue. The ROIs were drawn on an ADC-map, generated from b = 0 mm^{-2}s and 600 mm^{-2}s, under the guidance of CE T1W GE images. The mean value of the voxel-wise fits within the ROIs was calculated.

Table 8.1. Sequence parameters.

<table>
<thead>
<tr>
<th></th>
<th>DWI</th>
<th>T1W GE</th>
</tr>
</thead>
<tbody>
<tr>
<td>FOV (RL × AP) (mm²)</td>
<td>432 × 108</td>
<td>400 × 353</td>
</tr>
<tr>
<td>Acquisition matrix</td>
<td>144 × 34</td>
<td>236 × 208</td>
</tr>
<tr>
<td>Slices</td>
<td>18</td>
<td>56</td>
</tr>
<tr>
<td>Slice thickness/gap (mm)</td>
<td>3.7/0.3</td>
<td>1.7/—</td>
</tr>
<tr>
<td>TR/TE/ΔTE (ms)</td>
<td>&gt; 2200/45/—</td>
<td>4.7/1.15/1.0</td>
</tr>
<tr>
<td>FA (°)</td>
<td>90</td>
<td>10</td>
</tr>
<tr>
<td>BW (Hz/voxel)</td>
<td>59 (phase direction)</td>
<td>1602 (frequency)</td>
</tr>
<tr>
<td>Parallel imaging</td>
<td>1.3 (AP)</td>
<td>2/1.5 (RL/AP)</td>
</tr>
<tr>
<td>Partial Fourier</td>
<td>0.8</td>
<td>no</td>
</tr>
<tr>
<td>Respiratory compensation</td>
<td>Respiratory trigger (navigator)</td>
<td>1 breath-hold</td>
</tr>
<tr>
<td>Fat saturation</td>
<td>Gradient reversal during slice Dixon reconstruction selection + SPIR</td>
<td></td>
</tr>
<tr>
<td>b-values (mm⁻²s) and</td>
<td>0 (15), 10 (9), 20 (9), 30 (9), 40 (9),</td>
<td></td>
</tr>
<tr>
<td>directions/averages between</td>
<td>50 (9), 75 (4), 100 (12), 150 (4), 250 (4),</td>
<td></td>
</tr>
<tr>
<td>brackets</td>
<td>400 (4) and 600 (16)</td>
<td></td>
</tr>
</tbody>
</table>

¹TR of the DWI acquisition was determined by the respiratory trigger interval, but it was at least 2200 ms. Abbreviations: DWI = diffusion-weighted imaging; FOV = Field of view; RL = right–left; AP = anterior–posterior; TR = repetition time; TE = echo time; ΔTE = increase in echo time; FA = flip angle; BW = bandwidth per voxel; SPIR = spectral presaturation with inversion recovery.
Comparison of methods
We quantitatively evaluated the performance of fitting algorithms and models considering the following four factors: the goodness of fit, the uniqueness of the fit parameters, the precision, and the fit parameters’ contrast between tumor and non-tumorous pancreatic tissue.

Goodness of fit
For all models fitted with the Levenberg-Marquardt least squares fit, we took the mean adjusted $R^2$ from both ROIs as a measure for goodness of fit.

Uniqueness
For the IVIM fit algorithms and the multi-parametric DWI model, we used a Spearman’s rank correlation test between the fit parameters to examine the unique nature of the fit parameters (significance level $\alpha = 0.05$). Fit parameter combinations with significant Spearman’s rho indicate that measuring both parameters has limited added value. For this purpose, the fit parameters were averaged over the three acquisitions per patient.

Precision
From the repeated measures we calculated the inter- and intra-session within-subject coefficient of variation (wCV) of the tumor ROI as a measure of precision [29].

Tumor contrast
We tested whether parameter values from the tumor ROI (averaged over the three acquisitions per patient) were different from the values in non-tumorous pancreatic tissue (Wilcoxon signed-rank test; significance level $\alpha = 0.05$). Also, we calculated the contrast between tumor and non-tumorous pancreatic tissue. Per patient, the contrast was defined as the difference in parameter value of tumor and non-tumorous pancreatic tissue divided by the mean parameter value of both tissues multiplied by 100%. Finally, we plotted precision (wCV) as a function of contrast.

Statistical analyses
A Wilcoxon signed-rank test over the patient pairs was performed to determine which models had significantly lower adjusted $R^2$ compared to the model with highest adjusted $R^2$, both for tumor and non-tumorous pancreatic tissue (significance level $\alpha = 0.05$). For wCV and contrast, tests were performed to compare the results between the different IVIM fit algorithms as well as between the different DWI models (all Levenberg-Marquardt least squares fits, including IVIM-free and IVIM-fixed). For the
IVIM fit algorithms, tests were performed per parameter \((D, f, D^*)\). For the DWI models, parameters were split into two groups: diffusion-related parameters \((D, ADC, ADC_{slow}, DDC, M_{ADC}, D_{app})\) and perfusion-related parameters \((f, f_1, f_2, D^*, \alpha, \sigma_{ADC}, K_{app})\). Tests were done between the best parameter of a group and the remaining parameters. To test whether parameters had higher \(wCV\) than the best parameter, a Wilcoxon signed-rank test was performed over the squared differences of the repeated measure \((m_1\) and \(m_2\)), divided by the squared mean \((\mu)\) of the population for that parameter: \((m_1 - m_2)^2/\mu^2\) (significance level \(\alpha = 0.05\)). To test whether parameters had lower contrast than the best parameter, a Wilcoxon signed-rank test was performed over the contrast per patient (significance level \(\alpha = 0.05\)). Due to the exploratory nature of this research, no post hoc corrections were applied to the statistical tests presented.

**Results**

In two out of forty-two acquisitions it was not possible to delineate the tumor. Therefore, intra-session \(wCVs\) were determined using twelve patients. In one patient, no non-tumorous pancreatic tissue was present in the images. Therefore, tumor contrast was only based on thirteen patients. The mean mask sizes were 7.6 cm\(^3\) (= 210 voxels, range 3.0–23.5 cm\(^3\)) for the tumor ROIs and 4.2 cm\(^3\) (= 118 voxels, range 1.3–8.1 cm\(^3\)) for the non-tumorous pancreatic tissue ROIs.

![Figure 8.1](image.png)

*Figure 8.1.* The normalized mean signal, averaged over all acquisitions, plotted as a function of the full range of b-values *left* and zoomed at low b-values *right* for signal from tumor (red) and pancreatic (green) ROIs. The lines indicate monoexponential fits over range \(b = 150-600\) mm\(^2\)/s (dotted), \(b = 10-150\) mm\(^2\)/s (striped) and, only for pancreatic tissue, \(b = 0-10\) mm\(^2\)/s (solid). The figure shows that the classical IVIM-model does not fully describe the DWI data, in particular in pancreatic tissue.
Figure 8.1 highlights that the two exponents of the IVIM-model are insufficient to fully describe the DWI data, in particular in non-tumorous tissue, and underlines the importance of comparing several competing models. Parameter maps were generated for all fit algorithms and DWI models and shown in Fig. 8.2.

**Figure 8.2.** Axial parameter maps of the fit parameters of two IVIM model fit algorithms and seven DWI models in a 74 year-old female with pancreatic adenocarcinoma. ROIs containing diffusely growing pancreatic tumor infiltrating the adrenal gland (red) and pancreatic tissue in the pancreatic tail (green) are shown. The CE T1W GE is added as a reference.
**Goodness of fit**

The goodness of fit in tumor tissue was highest for the stretched-exp model (adjusted $R^2 = 0.96$; Fig. 8.3). However, all models had adjusted $R^2 > 0.93$ in tumor tissue. In non-tumorous pancreatic tissue the IVIM-free model had highest adjusted $R^2$ (0.94). Mono-exp, Gaussian and Kurtosis models had significantly lower adjusted $R^2$ in tumor as well as non-tumorous tissue and, in particular, they had adjusted $R^2 < 0.90$ in non-tumorous pancreatic tissue.

![Figure 8.3. The mean adjusted $R^2$ in tumor (red) and pancreatic (green) tissue as a measure of goodness of fit. Bars indicate the standard deviation between patients. Lines indicate significant differences (*: $p < 0.05$, **: $p < 0.01$) between the model with highest adjusted $R^2$ in tumor (red) and pancreatic (green) tissue and the connected model. The p-values are given in Table 8.D from the Appendix IV.](image)

![Figure 8.4. Plots of the wCV as a function of contrast between tumor and pancreatic tissue for the diffusion-related parameter (purple) and other fit parameters (orange, teal).](image)
CHAPTER 8

Uniqueness
Except for $D & D^*$ and $f & D^*$ of IVIM-adaptive, IVIM-Bayesian-log and IVIM-Bayesian-lin, there were no significant correlations between fit parameters (Table 8.2). The weakest correlations were found in the tri-exp model.

Precision
The inter-session $wCV$s were on average 30% larger than the intra-session $wCV$s (Table 8.3 and Table 8.C from Appendix III), indicating a larger test-retest variation when scans are repeated on separate days than in the same session. From the IVIM fit algorithms, IVIM-Bayesian-lin had most repeatable $f$ and $D^*$, and the $wCV$ for $D$ was not significantly worse than the best $wCV$ for $D$ (IVIM-fixed). IVIM-fixed had most repeatable $D$, and its $f$ was not significantly worse than $f$ of the IVIM-Bayesian-lin. From the DWI models, $D$ of IVIM-fixed had overall best $wCV$ (6.7%) from the diffusion-related parameters and was significantly lower than $D$ of IVIM-free, $M_{ADC}$ and $D_{app}$. From the perfusion-related parameters $\alpha$ of the stretched-exp model was lowest (14.4%) and significantly better than all other perfusion-related parameters. The stretched-exp model was the only model that produced good $wCV$ (< 15%) for both parameters.

Table 8.2. Uniqueness.

<table>
<thead>
<tr>
<th></th>
<th>Spearman’s rho</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$D &amp; f$</td>
</tr>
<tr>
<td>IVIM-free</td>
<td>-0.09</td>
</tr>
<tr>
<td>IVIM-adaptive</td>
<td>0.12</td>
</tr>
<tr>
<td>IVIM-Bayesian-log</td>
<td>-0.48</td>
</tr>
<tr>
<td>IVIM-Bayesian-lin</td>
<td>0.45</td>
</tr>
<tr>
<td>IVIM-MLE</td>
<td>0.10</td>
</tr>
<tr>
<td>IVIM-fixed</td>
<td>0.22</td>
</tr>
<tr>
<td>$ADC_{slow}$ &amp; $ADC_{fast}$</td>
<td>0.42</td>
</tr>
<tr>
<td>Mono-exp-2</td>
<td>D &amp; $f_1$</td>
</tr>
<tr>
<td>Tri-exp</td>
<td>0.08</td>
</tr>
<tr>
<td>Stretched-exp</td>
<td>-0.36</td>
</tr>
<tr>
<td>$M_{ADC}$ &amp; $\sigma_{ADC}$</td>
<td>-0.36</td>
</tr>
<tr>
<td>Gaussian</td>
<td>$D_{app}$ &amp; $K_{app}$</td>
</tr>
</tbody>
</table>

* = significant correlation ($p < 0.05$)
### Table 8.3. Precision of parameters and contrast.

<table>
<thead>
<tr>
<th>Fit algorithms</th>
<th>Inter-session wCV</th>
<th>Contrast (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>D</td>
<td>f</td>
</tr>
<tr>
<td>IVIM-free</td>
<td>10.0*</td>
<td>40.9</td>
</tr>
<tr>
<td>IVIM-adaptive</td>
<td>8.5*</td>
<td>34.4</td>
</tr>
<tr>
<td>IVIM-Bayesian-log</td>
<td>12.6</td>
<td>52.2</td>
</tr>
<tr>
<td>IVIM-Bayesian-lin</td>
<td>7.2</td>
<td>25.7</td>
</tr>
<tr>
<td>IVIM-MLE</td>
<td>8.4*</td>
<td>35.8</td>
</tr>
<tr>
<td>IVIM-fixed</td>
<td>6.7</td>
<td>28.7</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>DWI models (least squares fits)</th>
<th>Diffusion-related</th>
<th>Perfusion-related</th>
<th>Diffusion-related</th>
<th>Perfusion-related</th>
</tr>
</thead>
<tbody>
<tr>
<td>IVIM-free</td>
<td>10.0*</td>
<td>40.9**</td>
<td>50.5**</td>
<td>0.7</td>
</tr>
<tr>
<td>IVIM-fixed</td>
<td>6.7</td>
<td>28.7*</td>
<td>0.9</td>
<td>87.8**</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Mono-exp</th>
<th>ADC</th>
<th>ADC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>7.7</td>
<td>10.9</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Mono-exp-2</th>
<th>ADC&lt;sub&gt;slow&lt;/sub&gt;</th>
<th>ADC&lt;sub&gt;fast&lt;/sub&gt;</th>
<th>ADC&lt;sub&gt;slow&lt;/sub&gt;</th>
<th>ADC&lt;sub&gt;fast&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>8.83</td>
<td>20.61*</td>
<td>2.0**</td>
<td>45.4**</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Tri-exp</th>
<th>D</th>
<th>f₁</th>
<th>f₂</th>
<th>D</th>
<th>f₁</th>
<th>f₂</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>7.6</td>
<td>63.2*</td>
<td>36.7*</td>
<td>1.4**</td>
<td>36.7**</td>
<td>105.8</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Stretched-exp</th>
<th>DDC</th>
<th>α</th>
<th>DDC</th>
<th>α</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>8.0</td>
<td>14.4</td>
<td>13.9**</td>
<td>22.0**</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gaussian</th>
<th>M&lt;sub&gt;ADC&lt;/sub&gt;</th>
<th>σ&lt;sub&gt;ADC&lt;/sub&gt;</th>
<th>M&lt;sub&gt;ADC&lt;/sub&gt;</th>
<th>σ&lt;sub&gt;ADC&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>37.6**</td>
<td>35.6**</td>
<td>36.9</td>
<td>47.3**</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Kurtosis</th>
<th>D&lt;sub&gt;app&lt;/sub&gt;</th>
<th>K&lt;sub&gt;app&lt;/sub&gt;</th>
<th>D&lt;sub&gt;app&lt;/sub&gt;</th>
<th>K&lt;sub&gt;app&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>15.0*</td>
<td>36.9**</td>
<td>30.2*</td>
<td>24.6**</td>
</tr>
</tbody>
</table>

For both the diffusion-related parameter group as the perfusion-related parameter group, the parameter with lowest wCV is printed bold. Stars in indicate the parameters that were significantly (*: \( p < 0.05 \), **: \( p < 0.01 \)) worse than the best scoring parameter of the five groups (D, f, and D* for the algorithms; diffusion-related and perfusion-related for fit models). The p-values are reported in Table 8.E and 8.F from the Appendix IV. Abbreviations: wCV = within-subject coefficient of variation.

### Tumor contrast

Each fit method and algorithm, except for mono-exp, had at least one parameter that differed significantly between the tumor and non-tumorous pancreatic tissue (Table 8.4). From the fit algorithms, IVIM-Bayesian-lin was most promising, with a perfusion fraction that had significantly more contrast (contrast = 93.7%) than all
other algorithms (Table 8.3). From the DWI models, the greatest contrast (105.8%) was observed in $f_2$ of the tri-exp model, which was significantly larger than all other parameters (Table 8.3). Generally, parameters that showed a larger contrast between tumor and non-tumorous pancreatic tissue also had high wCVs (Fig. 8.4). However, the three parameters with the largest contrast ($f$: IVIM-Bayesian-lin; $f$: IVIM-fixed; $f_2$: tri-exp) had a relatively low wCV compared to their high contrast.

Finally, the results of all five tests are visually summarized in Fig. 8.5.

**Table 8.4(a).** Parameter values in tumor tissue.

<table>
<thead>
<tr>
<th>Tumor tissue</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>D</strong> (10⁻³ mm²s⁻¹)</td>
<td>$f$ (%)</td>
<td><strong>D</strong>* (10⁻³ mm²s⁻¹)</td>
</tr>
<tr>
<td>IVIM-free</td>
<td>1.38±0.14</td>
<td>4.98±1.04**</td>
</tr>
<tr>
<td>IVIM-adaptive</td>
<td>1.41±0.14</td>
<td>3.94±1.04**</td>
</tr>
<tr>
<td>IVIM-Bayesian-log</td>
<td>1.36±0.13</td>
<td>7.56±2.09*</td>
</tr>
<tr>
<td>IVIM-Bayesian-lin</td>
<td>1.41±0.14</td>
<td>2.56±0.81**</td>
</tr>
<tr>
<td>IVIM-MLE</td>
<td>1.43±0.15</td>
<td>4.42±0.99**</td>
</tr>
<tr>
<td>IVIM-fixed</td>
<td>1.50±0.15</td>
<td>2.60±0.75**</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>ADC</strong> (10⁻³ mm²s⁻¹)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mono-exp</td>
</tr>
<tr>
<td>Mono-exp-2</td>
</tr>
<tr>
<td>Tri-exp</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>DDC</strong> (10⁻³ mm²s⁻¹)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stretched-exp</td>
</tr>
<tr>
<td>Gaussian</td>
</tr>
</tbody>
</table>

| **D_app** (10⁻³ mm²s⁻¹) | $K_{app}$ |
|---------------------|
| Kurtosis | 2.17±0.26** | 0.91±0.16 |

Mean value±standard error (1.96 × SD/√n; $n = 14$ for tumor, $n = 12$ for pancreatic). * = significantly ($p < 0.05$) different from pancreatic tissue (Wilcoxon signed-rank test on individual patients). ** = significantly ($p < 0.005$) different from pancreatic tissue. The $p$-values are reported in Table 8.E from the supplemental materials III.
Table 8.4(b). Parameter values in pancreatic tissue.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>IVIM-free</th>
<th>IVIM-adaptive</th>
<th>IVIM-Bayesian-log</th>
<th>IVIM-Bayesian-lin</th>
<th>IVIM-MLE</th>
<th>IVIM-fixed</th>
</tr>
</thead>
<tbody>
<tr>
<td>$D$ (10^{-3} \text{mm}^2\text{s}^{-1})$</td>
<td>1.40±0.12</td>
<td>1.45±0.11</td>
<td>1.43±0.12</td>
<td>1.46±0.12</td>
<td>1.49±0.13</td>
<td>1.51±0.12</td>
</tr>
<tr>
<td>$f$ (%)</td>
<td>8.22±1.09</td>
<td>7.31±1.04</td>
<td>8.89±1.49</td>
<td>6.38±1.10</td>
<td>7.57±1.01</td>
<td>6.12±1.03</td>
</tr>
<tr>
<td>$D^*$ (10^{-3} \text{mm}^2\text{s}^{-1})$</td>
<td>93.6±21.3</td>
<td>98.1±21.3</td>
<td>246.3±111.1</td>
<td>101.2±13.4</td>
<td>99.2±21.4</td>
<td></td>
</tr>
</tbody>
</table>

Table 8.4(b). Parameter values in pancreatic tissue.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Mono-exp</th>
<th>Mono-exp-2</th>
<th>Tri-exp</th>
<th>Stretched-exp</th>
<th>Gaussian</th>
<th>Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\text{ADC}$ (10^{-3} \text{mm}^2\text{s}^{-1})$</td>
<td>1.75±0.19</td>
<td>1.44±0.13</td>
<td>1.41±0.12</td>
<td>1.83±0.23</td>
<td>0.59±0.14</td>
<td>2.83±0.30</td>
</tr>
<tr>
<td>$\text{ADC}_{\text{slow}}$ (10^{-3} \text{mm}^2\text{s}^{-1})$</td>
<td>$\text{ADC}_{\text{fast}}$ (10^{-3} \text{mm}^2\text{s}^{-1})$</td>
<td>3.05±0.29</td>
<td>0.74±0.06</td>
<td>$M_{\text{ADC}}$ (10^{-3} \text{mm}^2\text{s}^{-1})$</td>
<td>$\sigma_{\text{ADC}}$ (10^{-3} \text{mm}^2\text{s}^{-1})$</td>
<td>$D_{\text{app}}$ (10^{-3} \text{mm}^2\text{s}^{-1})$</td>
</tr>
<tr>
<td>$D$ (10^{-3} \text{mm}^2\text{s}^{-1})$</td>
<td>$f_1$ (%)</td>
<td>$f_1$ (%)</td>
<td>7.28±2.27</td>
<td>9.74±2.47</td>
<td>1.10±0.12</td>
<td></td>
</tr>
<tr>
<td>$\alpha$</td>
<td>$\text{DDC}$ (10^{-3} \text{mm}^2\text{s}^{-1})$</td>
<td>$\sigma_{\text{ADC}}$ (10^{-3} \text{mm}^2\text{s}^{-1})$</td>
<td>0.74±0.06</td>
<td>2.51±0.44</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Discussion

We evaluated the performance of the six most common IVIM model fit algorithms and seven DWI models in patients with pancreatic cancer by assessing the goodness of fit, the uniqueness and precision of the fit parameters, and the contrast between tumor and non-tumorous pancreatic tissue. Among the IVIM fit algorithms, the IVIM-Bayesian-lin and IVIM-fixed perform best as they exhibit the highest precision and best contrast for the perfusion parameter $f$. Among the models, the IVIM model scored relatively high in all tests, but for specific purposes (e.g. diagnostics and treatment response monitoring) other models may be preferred.
**IVIM fit algorithms**

From the fit algorithms tested, IVIM-Bayesian-lin performed best considering precision and tumor contrast. IVIM-Bayesian-lin showed the highest precision for $f$ and $D^*$ compared to the other algorithms, and the precision on $D$ was not significantly lower than for the algorithm with the highest precision (IVIM-fixed). Furthermore, the contrast between tumor and non-tumorous pancreatic tissue generated by $f$ was significantly higher than in all other algorithms. These findings are in agreement with earlier published simulations and volunteer measurements [10]. However, IVIM-Bayesian-lin is not widely implemented. Also, the uniqueness of fit parameters was poor for all three combinations of parameters in both Bayesian approaches. Therefore, IVIM-fixed may be a good alternative. IVIM-fixed showed the highest precision of $D$, and the precision in $f$ had second to best precision, which was not significantly worse than IVIM-Bayesian-lin. A disadvantage of IVIM-fixed is that no information on $D^*$ is obtained. However, all algorithms had strong correlations between $D^*$ and the other fit parameters, which implies a limited added value of $D^*$.

**DWI models**

The mono-exp, Gaussian, and kurtosis models overall scored poorly. This is explained by the fact that these models are unable to capture differences in the non-monoexponential behavior which separates tumour and non-tumorous pancreatic tissue. Therefore, these models had significantly lower adjusted $R^2$ compared to the DWI models with highest adjusted $R^2$, both in tumor and non-tumorous pancreatic tissue. As the kurtosis model and Gaussian model were not developed to describe DWI data from $b$-values $< 150 \text{mm}^2\text{s}$, it is interesting that they both had parameters that differed significantly between tumor and non-tumorous pancreatic tissue.

Overall, the IVIM model performed well, but for specific purposes, other models may be preferred. For diagnostics, when the goal is to distinguish between pancreatic tumors and non-tumorous pancreatic tissue, the model parameter with the highest contrast is preferred. We found that, in general, an increase in contrast came at the cost of a decrease in precision. However, for the parameters with highest contrast, the perfusion parameters of tri-exp ($f_2$), IVIM-fixed ($f$) and IVIM-Bayesian-lin ($f$), the increase in tumor contrast was to a much lesser extent accompanied by low precision. The $f_2$ parameter of tri-exp had significantly higher contrast than all other parameters, though the contrast of $f$ of IVIM-fixed and IVIM-Bayesian-lin was still less than 20% lower. Therefore, these parameters are best at distinguishing between pancreatic tumors and non-tumorous pancreatic tissue.
For treatment evaluation and response monitoring, two aspects must be taken into account. First, the relevant model parameter needs to be measured with high precision and hence a low wCV. Second, there should be a change in the parameter of interest as a result of the treatment. The parameters with lowest wCV was D of the IVIM-fixed model, which is a diffusion related parameter. Tissue diffusion has been reported as a good biomarker for treatment response for e.g. chemotherapy of colorectal hepatic metastasis [30]. The wCV of D of IVIM-fixed was not significantly lower than wCV of many other diffusion related parameters: D (IVIM-Bayesian-lin, IVIM-Bayesian-log, tri-exp), ADC (mono-exp), $ADC_{slow}$ (mono-exp-2) and DDC (stretched-exp). Consequently, several models can be used if diffusion is used as biomarker for treatment response. Perfusion-related parameters may be more sensitive to probe angiogenic changes as a result of therapy [7, 8]. In such a situation, $f$ (IVIM) would be the most straightforward
to use, as it directly relates to a perfusing blood fraction. Still, one should be aware of the limited precision of most perfusion-related parameters. The latter is not the case for the parameter $\alpha$ of the stretched-exp model. This parameter is related to the non-monoexponential behavior of the data as induced by perfusion. Furthermore, it had significantly lower wCV than all perfusion-related parameters. Unfortunately, there is no straightforward connection between $\alpha$ and physical perfusion parameters. Nevertheless, changes in $\alpha$ should relate to angiogenic changes and thus $\alpha$ may be a good candidate to monitor angiogenic changes. Hence, the stretched-exp model could be most useful for treatment response monitoring with two precise parameters ($DDC$ and $\alpha$) probing both diffusion and perfusion effects. This is confirmed by the findings by Orton et al. [31].

**Limitations**

A limitation of this study is that the ROI delineations were based on CE T1W GE and ADC-maps from $b=0$ and 600 mm$^2$s. Therefore, the ROI contained regions with low perfusion (CE T1W GE, ADC-map) and, potentially, diffusion restriction (ADC-map), which might have biased our perfusion measurements in tumors, compared to non-tumorous pancreatic tissues. However, so far, this is considered the best way to delineate pancreatic tumors.

We realize that with the inclusion of 14 patients, the accuracy of the estimation of the wCV is moderate (95% confidence interval is still 35% of the SD for $n=14$). We believe that for comparing the different algorithms and models this number suffices, especially due to the paired nature of the comparison.

In this study, we chose to assess multiple fit algorithms only for the IVIM model and not for the other models. This focus is justified by the fact that currently IVIM is the most widely used model for DWI data next to the monoexponential model. In addition, for the other DWI models alternative fit algorithms have not been widely used previously. Potentially, the other models may also improve to some extent if alternative fits are used.

**Conclusion**

When IVIM modeling of DWI data from pancreatic cancer patients is desired, there are two preferred options: IVIM-Bayesian-lin and IVIM-fixed. Overall, the IVIM model scored well but, depending on the application, it may be useful to also consider the triexponential model (promising for tumor detection) or stretched exponent model (treatment response monitoring).
References


Appendix I

DWI was acquired using respiratory triggering. All slices of the volume were acquired during one trigger to avoid inter-trigger mismatches between slices of one single volume (one b-value in one diffusion direction). This approach resulted in long acquisition periods each trigger (1.8 s). Patients were instructed to hold their breath during the typical noise produced by the EPI readout and to breathe freely during the navigator acquisition to minimize intra-trigger respiratory motion.

All DWI images were denoised using a Rician adaptive non-local means filter [19], with a search radius of three voxels and a patch radius of one voxel. The images were registered using a two-step approach in Elastix [20]. First, a reference image was created by averaging at least five manually selected acquisitions from the same respiratory position. To correct for bulk displacements between acquisitions, we performed a mutual information based rigid Euler transformation on each b-value to this reference image. The second step used a non-rigid b-spline registration based on mutual information to adjust for further deformations. During this step, we used two registration approaches: a single group-wise 4D [21] registration, or multiple 3D registrations for each b-value. We manually selected the registration approach that resulted in the most stable anatomy across the images acquired at different b-values and gradient directions.

Appendix II

Details on the IVIM fit algorithms are described in Table 8.A. Details of the fit models and their constraints are described in Table 8.B. For the tri-exp model, the pseudo-diffusion coefficients $D^*_1$ and $D^*_2$ were fixed to 0.014 and 0.093 mm$^2$s$^{-1}$ (based on healthy pancreatic data in 16 volunteers, data not shown). Note that $S_0$ was fitted too, instead of fixed to value from the $b = 0$ mm$^2$s$^{-1}$ acquisition. However, results of $S_0$ were not analyzed and $S_0$ further.

Per b-value we averaged the data to increase fitting speed. The mean data from b-values obtained in 4, 9, 12, 15 and 16 directions were weighted 1, 2, 3, 4 and 4 times during the fit, respectively. All fits were done voxel-wise.

Voxels with IVIM model parameter $f > 25\%$ contained mainly large vessels with instantaneous signal decay, which greatly influences $f$ and $D^*$ in the IVIM model [9]. For this reason, we removed voxels with $f > 25\%$ when determining mean $f$ and $D^*$ from the IVIM algorithms. In addition, when $f < 1\%$, there is too little perfusion signal to determine $D^*$. Therefore, only signal from voxels with $f > 1\%$ are considered when calculating mean $D^*$ from ROIs.
We tested seven fit models in this study: the mono-exp, mono-exp-2, IVIM, tri-exp, stretched-exp, Gaussian and kurtosis model. The mono-exp model returned an apparent diffusion coefficient (ADC) value from a fit to all data, whereas the mono-exp-2 model returns an $ADC_{\text{slow}}$ by fitting mono-exp to data from $b = 100 \text{mm}^2\text{s}^{-1}$ and $b = 600 \text{mm}^2\text{s}^{-1}$ and an $ADC_{\text{fast}}$ by fitting mono-exp to data $b = 0 \text{mm}^2\text{s}^{-1}$ and $b = 100 \text{mm}^2\text{s}^{-1}$. Mono-exp-2 probes IVIM-like features as $ADC_{\text{slow}}$ represents diffusion and $ADC_{\text{fast}}$ is also sensitive to perfusion but does not require as long a measurement time as typically seen for IVIM modeling. The tri-exp model described data from three different compartments with different (pseudo-) diffusion speeds ($D_1$, $D^*$, and $D^*_2$). $D_1$ relates to capillary perfusion, and $D^*_2$ relates to instantaneous dephasing of spins. The stretched-exp model [15] as well as the Gaussian model [16], which were also fitted, describe a continuous distribution of diffusivities rather than discrete diffusion values within a voxel. The Gaussian model does not imply the assumption that molecules diffuse Gaussian, but assumes that within a voxel, there is a single Gaussian distribution of ADC-values. Finally, the kurtosis model allows to incorporates non-Gaussian behavior of restricted diffusing molecules [32]. All models had parameters constrained to an appropriate domain (Table 8.B).

**Table 8.A. Fit algorithms.**

<table>
<thead>
<tr>
<th>Name</th>
<th>Fit</th>
</tr>
</thead>
<tbody>
<tr>
<td>IVIM-free</td>
<td>Levenberg-Marquardt least squares</td>
</tr>
<tr>
<td>IVIM-adaptive</td>
<td>Adaptive threshold segmented fit</td>
</tr>
<tr>
<td>IVIM-Bayesian-log* [23, 24]</td>
<td>Data-driven Bayesian fit for which the prior is a fitted Gaussian in log-space to confine parameters to relevant values</td>
</tr>
<tr>
<td>IVIM-Bayesian-lin* [10]</td>
<td>Data-driven Bayesian fit using boxcar functions with support over pre-defined ranges as weakly informative priors.</td>
</tr>
<tr>
<td>IVIM-MLE [25]</td>
<td>Maximum likelihood estimator approach that assumed Rician noise</td>
</tr>
<tr>
<td>IVIM-fixed</td>
<td>Levenberg-Marquardt least squares, except that $D^*$ was fixed to $70 \times 10^{-3}\text{mm}^2\text{s}^{-1}$, which resulted in more stable fits in healthy volunteers [9] (value based on volunteer data).</td>
</tr>
</tbody>
</table>

* Both Bayesian fits required a region of interest (ROI) containing data used for the data-driven fit. To determine the prior, a ROI larger than typical tumor size was required and hence ROIs were selected containing the tumor, pancreas, and surrounding tissue.
Table 8.8. DWI models.

<table>
<thead>
<tr>
<th>Formula</th>
<th>Parameters</th>
<th>Constraints</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\frac{S(b)}{S_0} = \text{Exp}[-ADC \cdot b]$</td>
<td>ADC = apparent diffusion coefficient</td>
<td>—</td>
</tr>
<tr>
<td>$\frac{S(b)}{S_0} = (1-f) \cdot \text{Exp}[-D \cdot b] + \frac{f}{D^<em>} \cdot \text{Exp}[-(D+D^</em>) \cdot b]$</td>
<td>$D = \text{diffusion coefficient}$</td>
<td>$0.5 \times 10^{-3} - 6 \times 10^{-3} \text{ mm}^{-2}\text{s}^b$</td>
</tr>
<tr>
<td></td>
<td>$D^* = \text{pseudo diffusion coefficient}$</td>
<td>$6 \times 10^{-3} - 200 \times 10^{-3} \text{ mm}^{-2}\text{s}$</td>
</tr>
<tr>
<td></td>
<td>$f = \text{perfusion fraction}$</td>
<td>$0.1 - 99%$</td>
</tr>
<tr>
<td>$\frac{S(b)}{S_0} = (1-f_1-f_2) \cdot \text{Exp}[-D \cdot b] + f_1 \cdot \text{Exp}[-D_1^* \cdot b] + f_2 \cdot \text{Exp}[-D_2^* \cdot b]$</td>
<td>$D_1^* = \text{pseudo diffusion coefficient 1}$</td>
<td>$0 - 10 \times 10^{-3} \text{ mm}^{-2}\text{s}$</td>
</tr>
<tr>
<td></td>
<td>$D_2^* = \text{pseudo diffusion coefficient 2}$</td>
<td>$14 \times 10^{-3} \text{ mm}^{-2}\text{s}$</td>
</tr>
<tr>
<td></td>
<td>$f_1 = \text{tissue fraction with } D_1^*$</td>
<td>$0.1 - 70%$</td>
</tr>
<tr>
<td></td>
<td>$f_2 = \text{tissue fraction with } D_2^*$</td>
<td>$0.1 - 70%$</td>
</tr>
<tr>
<td>$\frac{S(b)}{S_0} = \text{Exp}[-(b \cdot DDC)^\alpha]$</td>
<td>$DDC = \text{distributed diffusion coefficient}$</td>
<td>$0 - 6 \times 10^{-3} \text{ mm}^{-2}\text{s}$</td>
</tr>
<tr>
<td></td>
<td>$\alpha = \text{stretching coefficient}$</td>
<td>$0 - 2$</td>
</tr>
<tr>
<td>$\frac{S(b)}{S_0} = \left(1 + \frac{M_{\text{ADC}}}{\sigma_{\text{ADC}} \cdot \sqrt{2}} \cdot \frac{b \cdot \sigma_{\text{ADC}}}{\sqrt{2}}\right)$</td>
<td>$M_{\text{ADC}} = \text{apparent diffusion coefficient distribution maxima}$</td>
<td>$0.01 - 10 \times 10^{-3} \text{ mm}^{-2}\text{s}$</td>
</tr>
<tr>
<td></td>
<td>$\sigma_{\text{ADC}} = \text{ADC distribution width}$</td>
<td>$0.1 - 100 \times 10^{-3} \text{ mm}^{-2}\text{s}$</td>
</tr>
<tr>
<td>$\frac{S(b)}{S_0} = \text{Exp}[-b \cdot M_{\text{ADC}} + \frac{1}{2} \sigma_{\text{ADC}}^2 b^2]$</td>
<td>$K_{\text{app}} = \text{apparent diffusion kurtosis}$</td>
<td>$0.5 \times 10^{-3} - 14 \times 10^{-3} \text{ mm}^{-2}\text{s}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$0 - 10$</td>
</tr>
</tbody>
</table>

$S(b)$ is the signal from an acquisition with $b$-value $b$, $S_0$ is the signal with no diffusion-weighting and $\phi$ is the Gaussian error function. * This is the simple IVIM formula; we implemented the $TR$, $TE$, $T_1$ and $T_2$ corrected version [6]. * $D$ had no constraints in the IVIM-adaptive approach; IVIM-Bayesian-log had the following constraints: $D > 0 \text{ mm}^{-2}\text{s}$, $D^* > 0 \text{ mm}^{-2}\text{s}$ and $0 \% < f < 100\%$. 
## Appendix III

### Table 8.C. Intra-session wCV.

<table>
<thead>
<tr>
<th>Fit algorithms</th>
<th>$D$</th>
<th>$f$</th>
<th>$D^*$</th>
</tr>
</thead>
<tbody>
<tr>
<td>IVIM-free</td>
<td>6.3</td>
<td>32.9</td>
<td>32.7</td>
</tr>
<tr>
<td>IVIM-adaptive</td>
<td>5.3</td>
<td>33.9</td>
<td>39.3</td>
</tr>
<tr>
<td>IVIM-Bayesian-log</td>
<td>9.8</td>
<td>42.7</td>
<td>112.0</td>
</tr>
<tr>
<td>IVIM-Bayesian-lin</td>
<td>5.8</td>
<td>27.7</td>
<td>26.3</td>
</tr>
<tr>
<td>IVIM-MLE</td>
<td>5.0</td>
<td>38.7</td>
<td>41.8</td>
</tr>
<tr>
<td>IVIM-fixed</td>
<td>5.0</td>
<td>23.8</td>
<td></td>
</tr>
<tr>
<td><strong>ADC</strong></td>
<td><strong>4.5</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mono-exp</td>
<td>$\text{ADC}_{\text{slow}}$</td>
<td>$\text{ADC}_{\text{fast}}$</td>
<td></td>
</tr>
<tr>
<td>Mono-exp-2</td>
<td>6.32</td>
<td>18.04</td>
<td></td>
</tr>
<tr>
<td>Tri-exp</td>
<td>4.8</td>
<td>47.7</td>
<td>37.0</td>
</tr>
<tr>
<td>Stretched-exp</td>
<td>3.8</td>
<td>6.9</td>
<td></td>
</tr>
<tr>
<td>Gaussian</td>
<td>$M_{\text{ADC}}$</td>
<td>$\sigma_{\text{ADC}}$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>27.4</td>
<td>23.9</td>
<td></td>
</tr>
<tr>
<td>Kurtosis</td>
<td>$D_{\text{app}}$</td>
<td>$K_{\text{app}}$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>7.1</td>
<td>24.2</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: wCV = between-subject coefficient of variation
## Appendix IV

**Table 8.D.** The p-values (Wilcoxon signed-rank) for the adjusted $R^2$ compared to the best.

<table>
<thead>
<tr>
<th></th>
<th>Adjusted $R^2$</th>
<th>Pancreatic</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Tumor</td>
<td></td>
</tr>
<tr>
<td>IVIM-free</td>
<td>0.017</td>
<td>Best</td>
</tr>
<tr>
<td>IVIM-fixed</td>
<td>0.946</td>
<td>0.173</td>
</tr>
<tr>
<td>Mono-exp</td>
<td>&lt; 0.001</td>
<td>0.005</td>
</tr>
<tr>
<td>Tri-exp</td>
<td>0.168</td>
<td>0.091</td>
</tr>
<tr>
<td>Stretched-exp</td>
<td>Best</td>
<td>0.153</td>
</tr>
<tr>
<td>Gaussian</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>&lt; 0.001</td>
<td>0.042</td>
</tr>
</tbody>
</table>
### Table 8.E. The p-values for the inter-session wCV.

<table>
<thead>
<tr>
<th>Fit algorithms</th>
<th>Inter-session wCV</th>
<th>Diffusion-related</th>
<th>Perfusion-related</th>
</tr>
</thead>
<tbody>
<tr>
<td>IVIM-free</td>
<td>0.030</td>
<td>0.463</td>
<td>0.007</td>
</tr>
<tr>
<td>IVIM-adaptive</td>
<td>0.020</td>
<td>0.104</td>
<td>0.017</td>
</tr>
<tr>
<td>IVIM-Bayesian-log</td>
<td>0.104</td>
<td>0.035</td>
<td>0.035</td>
</tr>
<tr>
<td>IVIM-Bayesian-lin</td>
<td>0.153</td>
<td>Best</td>
<td>Best</td>
</tr>
<tr>
<td>IVIM-MLE</td>
<td>0.042</td>
<td>0.268</td>
<td>0.011</td>
</tr>
<tr>
<td>IVIM-fixed</td>
<td>Best</td>
<td>0.761</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>DWI models (least squares fit)</th>
<th>Inter-session wCV</th>
<th>Diffusion-related</th>
<th>Perfusion-related</th>
</tr>
</thead>
<tbody>
<tr>
<td>IVIM-free</td>
<td>0.030</td>
<td>0.007</td>
<td>0.007</td>
</tr>
<tr>
<td>IVIM-fixed</td>
<td>Best</td>
<td>0.035</td>
<td></td>
</tr>
</tbody>
</table>

| ADC Mono-exp                   | 0.325             |                   |                   |
| ADC Mono-exp-2                 | ADC\textsubscript{slow} | ADC\textsubscript{fast} | 0.135             | 0.042             |
| Tri-exp                        | 0.241             | 0.001             | 0.049             |
| Stretched-exp                  | 0.194             | Best              |                   |
| Gaussian                       | < 0.001           | < 0.001           |                   |
| Kurtosis                       | 0.013             | < 0.001           |                   |

Data are split into 5 groups (D, f and D* for the algorithms; diffusion-related and perfusion-related for fit models). The p-values of the Wilcoxon signed-rank test between the parameter with lowest wCV (Best) of each group and the rest of that group are noted. Abbreviations: wCV = between-subject coefficient of variation.
### Table 8.F. The p-values for the contrast parameters.

<table>
<thead>
<tr>
<th>Fit algorithms</th>
<th>Tumor and pancreatic</th>
<th>Contrast</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$D$</td>
<td>$f$</td>
</tr>
<tr>
<td>IVIM-free</td>
<td>0.424</td>
<td>0.005</td>
</tr>
<tr>
<td>IVIM-adaptive</td>
<td>0.233</td>
<td>0.016</td>
</tr>
<tr>
<td>IVIM-Bayesian-log</td>
<td>0.110</td>
<td>0.021</td>
</tr>
<tr>
<td>IVIM-Bayesian-lin</td>
<td>0.151</td>
<td>0.003</td>
</tr>
<tr>
<td>IVIM-MLE</td>
<td>0.339</td>
<td>0.016</td>
</tr>
<tr>
<td>IVIM-fixed</td>
<td>0.380</td>
<td>0.007</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ADC models (least square fit)</th>
<th>Diffusion-related</th>
<th>Perfusion-related</th>
</tr>
</thead>
<tbody>
<tr>
<td>IVIM-MLE</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>IVIM-fixed</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Mono-exp</th>
<th>ADC</th>
<th>ADC</th>
</tr>
</thead>
<tbody>
<tr>
<td>D</td>
<td>f</td>
<td></td>
</tr>
<tr>
<td>Mono-exp-2</td>
<td>$ADC_{slow}$</td>
<td>$ADC_{fast}$</td>
</tr>
<tr>
<td>D</td>
<td>$f_1$</td>
<td>$f_2$</td>
</tr>
<tr>
<td>Tri-exp</td>
<td>D</td>
<td>$f_1$</td>
</tr>
<tr>
<td>DDC</td>
<td>$\alpha$</td>
<td>$DDC$</td>
</tr>
<tr>
<td>Stretched-exp</td>
<td>$M_{ADC}$</td>
<td>$\sigma_{ADC}$</td>
</tr>
<tr>
<td>Gaussian</td>
<td>$D_{app}$</td>
<td>$K_{app}$</td>
</tr>
</tbody>
</table>

Left: p-value for Wilcoxon signed-rank test between parameter values in tumor and pancreatic tissue. Right: data are split into 5 groups ($D$, $f$, and $D^*$ for the algorithms; diffusion-related and perfusion-related for fit models). The p-value for Wilcoxon signed-rank test comparing contrasts from the parameter with the highest contrast (Best), compared to the remaining parameters per group are noted.