MRI of pancreatic cancer for radiotherapy

Gurney-Champion, O.J.

Citation for published version (APA):
CHAPTER 7

Minimizing the acquisition time for IVIM model MRI

Minimizing the acquisition time for intravoxel incoherent motion magnetic resonance imaging acquisitions in the liver and pancreas

Oliver J Gurney-Champion
Martijn Froeling
Remy Klaassen
Jurgen H Runge
Arjan Bel
Hanneke WM van Laarhoven
Jaap Stoker
Aart J Nederveen

Investigative Radiology 2016;51(4):211–220
Abstract

Objective: The aim of this study was to determine the combination of b-values and signal averages for diffusion-weighted image acquisitions that render the minimum acquisition time necessary to obtain values of the intravoxel incoherent motion (IVIM) model parameters in vivo in the pancreas or liver with acceptable reproducibility.

Materials and Methods: For 16 volunteers, diffusion-weighted images, with 14 b-values and 9 acquisitions per b-value, were acquired in 2 scan sessions. The IVIM model was fitted to data from lesion-sized regions of interest (ROIs) (1.7 cm$^3$) as well as organ-sized ROIs in the pancreas and liver. By deleting data during analyzes, the IVIM model parameters, $D$ and $f$, could be determined as a function of the number of b-values as well as the number of measurements per b-value taken along. For the IVIM model parameters, we examined the behavior reproducibility, in the form of the within-subject coefficient of variation ($wCV$), as a function of the amount of data taken along in the fits. Finally, we determined the minimum acquisition time required as a function of $wCV$.

Result: For the lesion-sized ROI, the intersession $wCVs$ were 8%/46% and 13%/55% for $D/f$ in the pancreas and liver, respectively, when all data were taken along. For 1.2 times larger $wCVs$, acquisition in the pancreas could be done in 5:15 minutes using 9 acquisitions per b-value at $b = 0, 30, 50, 65, 100, 375$ and 500 mm$^2$s and for the liver in 2:15 using 9 acquisitions per b-value at $b = 0, 40$ and 500 mm$^2$s.

Conclusions: Acquiring 7 b-values in the pancreas and 3 b-values in the liver only decreases the reproducibility by 20% compared with an acquisition with 14 b-values. The understanding of the behavior of reproducibility as a function of b-values and acquisitions per b-values scanned will help researchers select the shortest IVIM protocol.
Introduction

Diffusion-weighted imaging (DWI) is a promising technique for imaging pancreas and liver diffusivity and diffusivity-based lesion identification and characterization in these organs [1-3]. In DWI, the magnetic resonance (MR) signal is made sensitive to the diffusion of water molecules before the acquisition of the signal. In the classical diffusion model, the signal from diffusing water molecules is assumed Gaussian. Therefore, the signal attenuation is modeled monoexponentially, as a function of the apparent diffusion coefficient (ADC) of the tissue, and the amount of diffusion weighting of the signal (b-value).

However, DWI signal also attenuates due to capillary perfusion. Therefore, the intravoxel incoherent motion (IVIM) model for DWI, a biexponential model, was introduced [4]. This model uses three tissue-specific IVIM parameters to describe the signal attenuation as a function of b-value: the diffusion coefficient $D$, the pseudo-diffusion coefficient $D^*$, and the perfusion fraction $f$. In the pancreas and liver, data obtained with $b$-values higher than 150 mm$^{-2}$s are more predominantly sensitive to diffusion ($D$), whereas data obtained with $b$-values lower than 150 mm$^{-2}$s are also sensitive to perfusion ($D + D^*$). Since its introduction, the biexponential signal attenuation was confirmed in multiple studies [5-19] and related to perfusion [20, 21]. The added perfusion parameters of the IVIM model have shown additional value for lesion characterization in the pancreas and liver [5-11, 17, 22, 23]. In addition, they can be used to evaluate organ perfusion and diffusivity and discriminate between healthy liver and liver fibrosis [12-14]. They also enable treatment response monitoring in various other organs [15, 16]. Finally, $f$ shows an excellent histological correlation with histological tumor features in pancreatic tumors [24].

Determining IVIM parameters requires DW images acquired at multiple $b$-values, resulting in prolonged measurement time compared with the standard monoexponential model. For example, in studies that used respiratory compensation, the mean acquisition time of the pancreas and liver acquisitions was 8:30 minutes (range, 2:24–12:00 minutes; median, 7:55 minutes) [5, 6, 11, 13, 14, 22, 25, 26]. Currently, there is no criterion standard for the choice in $b$-values at which images are acquired and the number of images per $b$-value that should be acquired. The lack of a criterion standard could lead to either unnecessarily long acquisitions or biased and unrepeatable IVIM model parameter fits. Therefore, to have IVIM acquisitions implemented clinically, more insight should be available on the relation between the $b$-value selection and the number of averages, and the reproducibility and bias of the fitted IVIM model parameters.

Lemke et al. [27] used Monte-Carlo simulations to determine the ideal $b$-value distribution for a range of IVIM parameter values. They showed that at least 10
b-values need to be acquired for IVIM model to retrieve the parameters used in the simulated data. However, in simulations, the anatomy is assumed to be static. This situation does not hold in practical pancreatic and liver imaging because these organs are highly mobile due to respiratory, peristaltic, and cardiac motion. This motion can lead to misaligned neighboring slices and misalignments between different volumes belonging to different or repeated b-values. In addition, the echo planar imaging (EPI) readout used in DWI is prone to geometric distortions and susceptibility artifacts [28]. These artifacts often occur in the vicinity of air-tissue boundaries at the edge of the lungs, the stomach or the intestines. Finally, the IVIM model is only an approximation of the actual signal attenuation [4, 29] and the measured signal attenuation is influenced by other, additional factors than the IVIM model accounts for [30, 31]. This may lead a bias in the outcome that depends on the b-values sampled (the equivalence of ADC values that are determined from $b = 0$ and 1000 mm$^2$/s being overestimated compared with ADC values that are determined from $b = 150$ and 1000 mm$^2$/s in DWI-data of a well-perfused organ) [32].

Together, the effect discussed previously may lead to a different reproducibility of and bias in the IVIM model parameters in vivo than expected from simulations that use biexponential data. Therefore, the purpose of this study was to determine the combination of b-values and DW images per b-value that renders the minimum acquisition time necessary to obtain reproducible values of the IVIM parameters without a large bias in the pancreas or liver, using in vivo data.

Methods
We included 16 healthy volunteers (8 male, 8 female, mean age 28 years old, range 19–47 years) who were scanned on a 3 T Ingenia (Philips Healthcare, Best, the Netherlands) MR imaging (MRI) scanner. The scanner had a maximum gradient strength of 45 mT/m and a peak slew rate of 200 mT/m per milisecond. Data were acquired with a 16-channel phased-array coil anterior to the volunteer and a 10-channel phased-array coil posterior to the volunteer. All volunteers gave written informed consent. Two MRI sessions (1 hour to 20 days apart; mean, 7 days) were performed. In the first MRI session, we scanned DWI twice to assess intrasession reproducibility. In the second session, we scanned DWI once, which we compared with the first scan of the first session to assess intersession reproducibility. For 2 volunteers, only intersession data were scanned. During each session, we also acquired a 3D Dixon (3 echoes) and a multislice 2D T2-weighted turbo spin echo (TSE) sequence as anatomical references (see Table 1 for acquisition details).
Table 7.1. Sequence parameters for the multi-slice 2D DWI (first column), the 3D Dixon (second column) and the multi-slice 2D $T_2$-weighted TSE (last column).

<table>
<thead>
<tr>
<th></th>
<th>DWI</th>
<th>Dixon</th>
<th>$T_2$-weighted TSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>FOV (mm$^2$)</td>
<td>$432 \times 108$ (RL × AP)</td>
<td>$400 \times 350$</td>
<td>$400 \times 369$</td>
</tr>
<tr>
<td>acquisition matrix</td>
<td>$144 \times 36$</td>
<td>$236 \times 208$</td>
<td>$308 \times 230$</td>
</tr>
<tr>
<td>Slices</td>
<td>21</td>
<td>53</td>
<td>45</td>
</tr>
<tr>
<td>Slice thickness/gap (mm)</td>
<td>$3.7/0.3$</td>
<td>$1.7/—$</td>
<td>$5/—$</td>
</tr>
<tr>
<td>TR/TE/ΔTE (ms)</td>
<td>$&gt; 2350/44/—$</td>
<td>$4.4/1.15/0.9$</td>
<td>$779/80/—$</td>
</tr>
<tr>
<td>FA (°)</td>
<td>90</td>
<td>10</td>
<td>90</td>
</tr>
<tr>
<td>BW (Hz/voxel)</td>
<td>62.5 (phase direction)</td>
<td>1602 (frequency)</td>
<td>548.4 (frequency)</td>
</tr>
<tr>
<td>SENSE</td>
<td>1.7 (AP)</td>
<td>2/1.5 (LR/AP)</td>
<td>2 (AP)</td>
</tr>
<tr>
<td>Respiratory compensation</td>
<td>Respiratory trigger (navigator)</td>
<td>1 breath-hold</td>
<td>3 breath-holds</td>
</tr>
<tr>
<td>Fat saturation</td>
<td>Gradient reversal during slice selection + SPIR</td>
<td>Dixon reconstruction</td>
<td>SPAIR</td>
</tr>
<tr>
<td>b-values (mm$^2$s)</td>
<td>0, 10, 20, 30, 40, 50, 65, 80, 100, 125, 175, 250, 375, 500</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

TR of the DWI acquisition was determined by the respiratory trigger interval, but at least 2350 ms. Abbreviations: FOV = Field of view; RL = right–left; AP = anterior–posterior; TR = repetition time; TE = echo time; FA = flip angle; BW = bandwidth; SPIR = spectral presaturation with inversion recovery; SPAIR = spectrally selective attenuated inversion recovery.

**DWI protocol**

We acquired all DW images with a single shot EPI readout with axial slice orientation (see Table 7.1 for acquisition details). Diffusion weighting was induced in 3 orthogonal directions, and 3 images were obtained per direction. This resulted in 9 acquisitions per b-value. Each of these acquisitions was acquired at a navigator based expiratory trigger that tracked the liver-lung transition. To avoid intertrigger mismatches between slices of 1 single volume, all slices in the volume were acquired during each trigger. This approach necessitates relatively long acquisition periods (2.35 seconds) compared with common respiratory acquisition times (< 1s). Therefore, to reduce intratrigger respiratory motion, the volunteers were instructed to hold their breath during the typical noise produced by the EPI readout and to breathe freely during the navigator acquisition. This instruction was given before the MRI exam using a recorder sound fragment of the sequence. The bandwidth (BW) per voxel in the phase encoding direction used in literature is on average 21 Hz/voxel (range 9–75 Hz) in the pancreas and liver [1, 5–9, 14, 18, 21, 22, 26, 33, 34]. We used a high BW per voxel (62.5 Hz/voxel)
in order to minimize geometric distortions and susceptibility artifacts. To achieve this, we used parallel imaging as well as a small field of view (FOV) in the phase encoding direction, combined with 2 saturation slabs placed to the posterior and anterior of the FOV.

**Image processing**

Delineations were done on an averaged $b = 0 \text{ mm}^2\text{s}$ image (averaged over the 9 acquisitions) under the guidance of an averaged $b = 100 \text{ mm}^2\text{s}$, water only reconstruction of the Dixon, and $T_2$-weighted TSE images. The first region of interest (ROI), $ROI_{\text{large}}$, consisted of the entire pancreas or 2 to 4 slices in the liver. The second ROI, $ROI_{\text{small}}$, was lesion sized ($1.7 \text{ cm}^3$) and spanned 3 slices in the pancreas tail or liver. The $ROI_{\text{small}}$s were mostly $3 \times 4 \times 4$ voxel squares located in 3 adjacent slices either in the center of the liver or the tail of the pancreas. However, the pancreas is a small organ and sometimes the squares were deformed to fit in the pancreas. Using anatomical landmarks, such as vessels, ducts, and organ edges, the ROIs were placed at the same location in the different scans for the same volunteer.

![Figure 7.1. Example images from an axial slice (c, d, g, h) and coronal multiplanar reformatted slice (a, b, e, f) of the data as acquired (a, c, e, g) and the data after denoising, cardiac artifact removal, registration, and averaging (b, d, f, h) for images from $b = 0 \text{ mm}^2\text{s}$ (a–d) and $b = 500 \text{ mm}^2\text{s}$ (e–h). An example of signal loss due to the cardiac artifact (i, blue circle) compared with an acquisition of the same slice and $b$-value without the signal loss (j). $ROI_{\text{large}}$ (green) from the liver on the $b = 0 \text{ mm}^2\text{s}$ (k) image as well as the ADC image (from $b = 0 \text{ mm}^2\text{s}$ and $10 \text{ mm}^2\text{s}$) (l) is shown. The parts selected by the vessel selection algorithm (red) were excluded from $ROI_{\text{large}}$.](image)
For data processing, a custom-made pipeline was developed in Matlab 2013a (MathWorks, Natick, MA) comprising the following steps. The DW images were denoised using an adaptive optimized non-local means Rician denoising filter as implemented by Manjón et al. [35], with a search radius of 3 voxels and a patch radius of 1 voxel. In some slices, local signal loss due to cardiac motion during diffusion encoding occurred (Fig. 7.1 i-j), which has been reported before [36]. To remove slices containing these cardiac artifacts, we wrote an algorithm that excluded these slices from further analyzes. For each slice and b-value, we calculated the mean and standard deviation (SD) of all voxel values within ROI\textsubscript{large} of the organ of interest, using all 9 measurements of the same slice. Slices that had a mean intensity in the organ of 2 SDs lower than the mean over all acquisitions of that slice and b-value were removed. On average, the algorithm deleted 7.3% (range between volunteers 3.7%–10.6%) of the slices for the pancreas and 7.9% (range, 4.8%–9.9%) for the liver. To correct for interrespiratory trigger geometric variations and the effect of eddy currents, we registered all DW images group wise using a 4D non-rigid b-spline algorithm based on mutual information using Elastix [37, 38]. By masking, we excluded the slices selected by the cardiac motion algorithm from the cost function of the registration. After registration, the delineations were optimized to correct for errors due to registration.

Per DWI set, the image processing was performed twice, once using the pancreas ROI and once using the liver ROI.

**Data fitting**

For the IVIM fits, we used a $T_1$ and $T_2$ compensated IVIM model [21] with $T_1 = 725$ milliseconds, $T_2 = 43$ milliseconds for the pancreas, $T_1 = 809$ milliseconds, $T_2 = 34$ milliseconds for the liver, as well as $T_1 = 1932$ milliseconds, $T_2 = 275$ milliseconds for blood [39, 40]. As repetition time ($TR$) was dependent on the respiratory trigger time, which fluctuated each respiratory cycle, we chose a $TR = 5000$ milliseconds for IVIM model fitting.

Signal coming from large blood vessels decays rapidly when diffusion gradients are being applied due to different coherent flow speeds within a voxel. This can cause the data to behave like a triexponential decay [41], rendering it inappropriate for the biexponential IVIM model. In the large ROIs, we segmented and removed such vessels from the ROI\textsubscript{large} using the DW images by excluding voxels with an ADC value higher than 0.022 mm$^2$s$^{-1}$ on an ADC image created from $b = 0$ and 10 mm$^{-2}$s [7]. The cutoff value of 0.022 mm$^2$s$^{-1}$ was chosen after visually comparing the performance of the algorithm for several cutoff values in the liver, as the vessels in the liver are easy to detect (typical example in Fig. 7.1 k-l). At this value, the algorithm segmented vessels
in the liver that were also visible on the $b = 0$ mm$^2$s DW image (Fig. 7.1) as well as the Dixon and T2-weighted images. To visualize the effect of the segmented vessels on the data from the ROIs, we plotted the data, averaged over all volunteers, from the ROI$_{\text{large}}$s before and after vessel segmentation. To check how well these data were described the IVIM model, we fitted a biexponential fit [4] and a triexponential fit [41] to it. For biexponential fits to the data from the ROIs with and without vessels we calculated the adjusted $R^2$ value to determine how well these fits described the data.

All fits were done using the unconstrained Levenberg-Marquardt least squares fit. Unless mentioned otherwise, we fixed the value of $D^*$ to the value obtained from the fits to the averaged data of all ROI$_{\text{large}}$s from all volunteers [5-7]. To examine the influence of fixing $D^*$ on $D$ and $f$, we determined $D$ and $f$ per patient per DWI acquisition, using both the $D^*$-free and $D^*$-fixed fitting methods. We compared for significant ($p \leq 0.05$) differences between both methods using a paired t-test for $D$ and $f$. In addition, to get an idea of fluctuations in $D$, $f$, and $D^*$ over this healthy population, we calculated the between-subject coefficient of variations (bCVs, also known as coefficient of variations), using the data from only the first scan of each subject.

### Reproducibility of parameter fits

To study the reproducibility and bias of IVIM model parameters as a function of the $b$-values and number of selected DW images per $b$-value in the fit, we repeatedly deleted data points before the IVIM model fit.

As a measure for reproducibility, we took the within-subject coefficient of variation (wCV) for $D$ and $f$: wCV$_D$ and wCV$_f$, respectively, for the different fits mentioned previously. The wCVs were calculated as defined by Barnhart et al. [42] (see Supplemental Digital Content 1, http://links.lww.com/RLI/A234, for a short description). We used a Levine test on the percentage difference to check for significant ($p < 0.05$) differences between the intersession and intrasession wCVs for the full data set (SPSS, version 21, IBM, New York, NY). In addition, for the full data set, Bland-Altman plots [43] were plotted for all ROIs to visualize intersession variations for the pancreas and liver. A 1-sample t-test was done to test for any significant bias ($p \leq 0.05$) of $D$ and $f$.

As a measure for bias, we took the normalized deviation (ND) of $D$ and $f$ from the $D$ and $f$ values obtained from the full data set. For example, for $D$, this was defined as

$$ND_D = \frac{D_{\text{mean}} - D_{\text{mean,full}}}{D_{\text{mean,full}}} \times 100\%$$

(7.1)
in which $D_{\text{mean}}$ was the value found for $D$ averaged over all volunteers using the specific combination of b-values and DW images per b-value and $D_{\text{mean,full}}$ was value found for $D$ averaged over all volunteers using the full data set. $ND_f$ was defined similarly.

To visualize the reproducibility and bias as a function of the data taken along, we plotted heat maps of $ND$ and $wCV$ as a function of the number of b-values and the acquisitions per b-value used. This was done for intersession and intrasession, liver and pancreas, ROI$_{\text{small}}$ and ROI$_{\text{large}}$ and with $D^*$ as a fixed and as a free parameter in the IVIM model. In addition, deleting b-values was done according to three different preselected schemes: a scheme in which mainly low b-values remain in the data set (b-low), a scheme in which mainly high b-values remain (b-high) and a scheme in which b-values are deleted randomly (b-both). In these schemes, b-values were removed in the following order: b-low: 175, 375, 65, 125, 80, 30, 250, 50, 10, 100, 20 and 40 mm$^{-2}$s, b-high: 20, 40, 10, 65, 175, 80, 30, 100, 375, 50, 250 and 125 mm$^{-2}$s and b-both: 125, 80, 40, 20, 250, 10, 175, 65, 375, 30, 50 and 100 mm$^{-2}$s. All schemes kept $b = 0$ and 500 mm$^{-2}$s. Images per b-value were deleted in a way that the directional spread of the diffusion weighting was as homogeneous as possible.

To investigate the relation between scan time and reproducibility, we selected the combinations of b-values, acquisitions per b-value and b-scheme that had both $wCV$s lower than 1 to 2 (steps of 0.1) times the $wCV$s from the full acquisitions. To ensure robustness of these combinations, we then selected the combinations of which the neighbors also satisfied the $wCV$ criterion. From these combinations, we selected the one with the lowest number of acquisitions (b-values $\times$ acquisitions per b-value). Finally, by assuming a respiratory cycle of 12 per minute, the shortest acquisition time was plotted as a function of increasing loss in reproducibility.

As the performance of our image processing may depend on the b-values and the number of selected DW images per b-value, we repeated the image processing for the combination that yielded the fastest acquisition without increasing the $wCV$s by more than 20% and compared the $wCV$s and $ND$s to the case where image processing was done with the full data set. In addition, $D$- and $f$-maps were calculated for these settings and compared with the $D$- and $f$-maps from the full data set.

Results

Data fitting

Using data from ROIs including large vessels, the IVIM model fit was inaccurate at describing data from low b-value when compared with a triexponential fit (Fig. 7.2).
When vessels were excluded, the IVIM model described the data as accurate as the triexponential fit, both for the pancreas and liver. The $R^2$ values of the biexponential fits were 0.91 and 0.77 for the pancreas and liver, respectively, when vessels were included and increased to 0.92 and 0.82 when vessels were excluded, which can also be visually appreciated in Fig. 7.2. Therefore, the ROI$_{large}$ excluding vessels was selected for further analyzes, which resulted in ROI$_{large}$ of 29 cm$^3$ on average (range, 18–43 cm$^3$) for the pancreas and 33 (range, 13–59) cm$^3$ for 2 to 4 slices of the liver. We found $D^* = 70.3 \times 10^{-3}$ mm$^2$s$^{-1}$ in the pancreas and $D^* = 58.6 \times 10^{-3}$ mm$^2$s$^{-1}$ in the liver. In ROI$_{small}$, the vessels were excluded during delineation.

$D$ and $f$ were calculated for the entire group using the $D^*$ fixed and $D^*$ free IVIM model fits (Table 7.2). There was no significant difference for the fitted $D$ between both fitting methods. However, $f$ was significantly different in all cases (liver, pancreas, ROI$_{small}$, and ROI$_{large}$). The $bCV$s (Table 7.3) were similar for the $D^*$-free and $D^*$-fixed fits.

Figure 7.2. Log plots of the data from the signal at different $b$-values averaged over all 3 measurements from all 16 volunteers of the pancreas (a and b) and the liver (c and d) using the large ROIs with (a and c) and without (b and d) large vessels. The IVIM model (green dashes) and a triexponential decay (red line) were fitted to the data. Subplots show a magnification of the low $b$-values. The error bars indicate the SD between volunteers of the first scan.
Reproducibility of parameter fits

Both intersession and intrasession wCVs were better for the pancreas than for the liver (Table 7.4). The wCVs were similar to the bCVs. Six out of 8 wCVs mentioned in Table 7.4 were lower for intrasession than intersession, and one was equal. However, none of the differences were significant according to the Levine test. The Bland-Altman plots (Fig. 7.3) from the ROIsmall for \( D \) and \( f \) did not show any significant bias.

The intersession and intrasession results on all individual b-schemes, all ROIs and both organs, for fits with a fixed \( D^* \) as well as \( D^* \) as free model parameter, were calculated (see graphs, Supplemental Digital Content 2, http://links.lww.com/RLI/A235, and Supplemental Digital Content 3, http://links.lww.com/RLI/A236, which show heat maps for all discussed situations). Generally, both for pancreas and liver, when data from fewer b-values and fewer DW images per b-value were selected, the wCVs and NDs became worse (e.g., Figs. 7.4 and 7.5). In addition, the heat maps changed between the different schemes for deleting b-values (Supplemental Digital Content 2, http://links.lww.com/RLI/A235, and Supplemental Digital Content 3, http://links.lww.com/RLI/A236).

Table 7.2. Mean values for the population ± the SD of the value over the population.

<table>
<thead>
<tr>
<th></th>
<th>( D^\text{*-fixed} )</th>
<th>( D^\text{*-free} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pancreas</td>
<td>( D \pm SD (10^{-3} \text{mm}^2\text{s}^{-1}) )</td>
<td>( f \pm SD )</td>
</tr>
<tr>
<td>ROI_large</td>
<td>1.41±0.17</td>
<td>5.4±1.1</td>
</tr>
<tr>
<td>ROI_small</td>
<td>1.34±0.18</td>
<td>6.9±2.6</td>
</tr>
<tr>
<td>Liver</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ROI_large</td>
<td>1.12±0.10</td>
<td>5.1±1.9</td>
</tr>
<tr>
<td>ROI_small</td>
<td>0.99±0.12</td>
<td>5.8±1.5</td>
</tr>
</tbody>
</table>

Table 7.3. Intersession and intrasession bCVs using 14 b-values and 9 measurements per b-value.

<table>
<thead>
<tr>
<th></th>
<th>( D^\text{*-fixed} )</th>
<th>( D^\text{*-free} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pancreas</td>
<td>( bCV_c ) (%)</td>
<td>( bCV_f ) (%)</td>
</tr>
<tr>
<td>ROI_large</td>
<td>13</td>
<td>22</td>
</tr>
<tr>
<td>ROI_small</td>
<td>15</td>
<td>52</td>
</tr>
<tr>
<td>Liver</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ROI_large</td>
<td>12</td>
<td>27</td>
</tr>
<tr>
<td>ROI_small</td>
<td>20</td>
<td>35</td>
</tr>
</tbody>
</table>
For the fixed $D^*$ model, the $|ND_D|$ was less than 5% and $|ND_f|$ was less than 15% for most of the combinations, meaning that decreasing the acquisition time did not introduce a large bias. When the number of acquisitions per b-value was 1, 2 or 4, $|ND|$ mostly increased. In these cases, one of the diffusion directions was strongly overrepresented, which was not the case in the reference data set (3 orthogonal acquisitions repeated 3 times). For more than 4 images per b-value, this effect decreased.

$D$ and $f$ from fits with fixed $D^*$ (e.g., Fig. 7.4) were more reproducible (lower $wCV$) and had less bias (lower $|ND|$) for shorter acquisitions than when $D^*$ was set as free fitting parameter (e.g., Fig. 7.5). In addition, when $D^*$ was set as free fitting parameter, the intersession $wCV$ of $D^*$ ($wCV_D$) for the full data set was large: 64% and 134% for ROI\textit{large} and ROI\textit{small}, respectively, in the pancreas and 150% and 77% for both ROIs in the liver. The intrasession $wCV$s for $D^*$ were 61% and 123% for ROI\textit{large} and ROI\textit{small}, respectively, in the pancreas and 133% and 88% for both ROIs in the liver.

When $wCV$s larger than the $wCV$s found from the full data set were accepted (Table 7.4), IVIM could be acquired in shorter scan times (Fig. 7.6). For the pancreas in ROI\textit{large} and the liver in ROI\textit{small}, shorter scan times could be used without decreasing the reproducibility. When $D^*$ was fixed, the scan time could be halved to quartered, without increasing the $wCV$ by more than 20% relative to the full acquisition, for all ROIs. When $D^*$ is set as a free fitting parameter, the decrease in scan time was less pronounced (see graphs, Supplemental Digital Content 3, http://links.lww.com/RLI/A236, Fig. 7 U, which is similar as Fig. 7.6 except $D^*$ is set as free fit parameter).

The fastest acquisition that had less than 20% additional error for ROI\textit{large} was a 3:30 minute scan for the pancreas comprising of 6 images at $b = 0$, 30, 50, 65, 100, 375 and 500 mm$^{-2}$s and a 2:15 minute scan in the liver comprising 9 images at $b = 0$, 125 and 500 mm$^{-2}$s. For ROI\textit{small}, reproducible acquisition was feasible in the pancreas in 5:15 minutes using 9 images at $b = 0$, 30, 50, 65, 100, 375 and 500 mm$^{-2}$s and in the liver in 2:15 minutes using 9 images at $b = 0$, 40 and 500 mm$^{-2}$s.

<table>
<thead>
<tr>
<th>ROIs</th>
<th>Intra</th>
<th>Inter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pancreas</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ROI\textit{large}</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td>ROI\textit{small}</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>Liver</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ROI\textit{large}</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>ROI\textit{small}</td>
<td>12</td>
<td>13</td>
</tr>
</tbody>
</table>

Table 7.4. Intersession and intrasession wCVs using 14 b-values and 9 measurements per b-value.
Figure 7.3. Bland-Altman plots for the intersession variation of $D$ and $f$ from ROI$_{small}$ of the pancreas (top) and liver (bottom) using all b-values and acquisitions per b-value.

Figure 7.4. From left to right, heat maps of intersession $wCV_D$, $wCV_f$, $ND_D$ and $ND_f$ for ROI$_{small}$ in the pancreas (top row) and liver (bottom row). The vertical axis shows the number of b-values selected, whereas the horizontal axis indicates the number of selected DW images per b-value. The plots show results for deleting b-values using the b-both scheme. White voxels indicate that no values were obtained from the fit, and thus no $wCV$ or $ND$ could be calculated for this combination.
Figure 7.5. Heat maps similar to the ones in Fig. 7.4 from the fits with $D^*$-free.

Figure 7.6. Plots of the scan times as a function of additional relative error on intersession $wCV$ for ROI\_{large} (left column) and ROI\_{small} (right column). The arrows indicate the settings at which Fig. 7.7 was produced.

When image post-processing was done only on the selected data from the fast acquisition of ROI\_{small}, the bias and reproducibility changed from $wCV_D = 9\%$, $wCV_f = 44\%$, $ND_D = 0.5\%$, $ND_f = 4.2\%$ to $wCV_D = 10\%$, $wCV_f = 46\%$, $ND_D = 0.7\%$, and $ND_f = 1.5\%$ in the pancreas and from $wCV_D = 14\%$, $wCV_f = 49\%$, $ND_D = 3.8\%$, $ND_f = 5.8\%$ to $wCV_D = 16\%$, $wCV_f = 51\%$, $ND_D = 4.9\%$, and $ND_f = 4.7\%$ in the liver. This indicates that our post-processing also performs well with fewer data. There was little difference in the $D$- and $f$-maps reconstructed from the full data set, compared with the limited data set (Fig. 7.7).
Discussion

We used in vivo DWI data to determine the minimal number of b-values and acquisitions per b-value necessary to obtain reproducible IVIM parameters in the pancreas and liver with low bias. First, we showed that the IVIM model is only valid in ROIs in which large vessels are excluded. Then, we showed that when fewer b-values and acquisitions per b-value are used, IVIM modeling is more reproducible with $D^*$ fixed than with $D^*$ set as a fitting parameter. Considering this and the poor wCV of $D^*$, we recommend that $D^*$ is fixed. In addition, we showed that for the pancreas IVIM can be done in 3:30 minutes when changes affecting the entire organ are of interest, and in 5:15 minutes when a specific lesion needs to be measured. Finally, for the liver IVIM can be obtained in 2:15 in both cases.

In this paper, we show that large vessels contribute to an extra attenuation component in the DW-data. This is because signals from flowing blood in large vessels are very sensitive to diffusion gradients, due to the parabolic velocity profile of the blood. Signal thus rapidly decays between $b = 0$ and $10 \text{ mm}^2\text{s}$ for such vessels. Therefore, we recommend that during post-processing, either large vessels are removed from the ROI, or the use of triexponential models is taken into consideration. Alternatively, flow-compensating gradients could be used to overcome this problem at the acquisition level [29].

![Figure 7.7](image)

**Figure 7.7.** An example of the performance of DWI after selecting optimal settings compared with the full scan. The **top row** shows the $D$- and $f$-maps obtained with the full set of data. The **middle row** shows the same maps when only the selected sub-set of data is taken along (post-processing performed after deleting b-values). In the **bottom row**, the $b = 0 \text{ mm}^2\text{s}$ is shown as a reference. The organ of interest (left images, pancreas; right images, liver) is delineated by arrows.
The $D^*$ found by fitting to all data was considerably lower than the mean $D^*$ from the fits to individual patients (Table 7.2). When all data are combined, the signal-to-noise (SNR) ratio becomes higher and $D^*$ can be estimated better than on individual level. In simulations, we indeed found that $D^*$ was overestimated when the SNR decreased (Supplemental Digital Content 4, http://links.lww.com/RLI/A237, simulations of fitted parameters as a function of SNR). The signs of the systematic errors found in these simulations were in agreement with the signs of the $N_D$s from the $D^*$-free fitting, in which $D^*$ and $f$ were overestimated and $D$ was underestimated (Fig. 7.5 and Supplemental Digital Content 3, http://links.lww.com/RLI/A236). In addition to a different $D^*$, there was a small, but significant, difference in $f$ between the $D^*$-fixed fits and $D^*$-free fits. This is a result of the different $D^*$ values used in latter.

The $D$ and $D^*$ values reported here for the pancreas and liver agreed well with recent literature: $D = 0.90$ to $1.48 \times 10^{-3}$ mm$^2$s$^{-1}$ and $D^* = 14.8$ to $59.4 \times 10^{-3}$ mm$^2$s$^{-1}$ in the pancreas [5, 9, 17, 21, 33] and $D = 1.00$ to $1.17 \times 10^{-3}$ mm$^2$s$^{-1}$ and $D^* = 13.1$ to $70.6 \times 10^{-3}$ mm$^2$s$^{-1}$ in the liver [13, 14, 17, 22, 23, 25]. Our fitted values for $f$ in the pancreas and liver are similar to the value found in the pancreas by Lemke et al. [21] ($f = 6.1\%–7.8\%$), who introduced the $T_1$ and $T_2$ correction that we also applied. To compare our values for $f$ with other literature, we recalculated the perfusion fraction, $f'$, as would be obtained from the classical IVIM model in an acquisition with typical values for $TR$ (4500 milliseconds) and echo time ($TE = 70$ milliseconds). We found an $f' = 16.9\%$ in the pancreas and an $f' = 22.8\%$ in the liver. This is similar to values found in healthy volunteers in other research: $f' = 16\%$ to $40\%$ for pancreas and $f' = 15\%$ to $32\%$ for liver [5, 9, 13, 14, 17, 21–23, 25, 33].

We found that when $D^*$ was fixed, acquisitions at only 7 (pancreas) or 3 (liver) b-values were required for reproducible IVIM modeling with low bias. This is lower than the minimum of 10 b-values required according to previous simulations [27]. However, those simulations also fit $D^*$, which requires additional data points. When $D^*$ is not fixed, we require 12 b-values for the pancreas and 9 b-values in the liver (see Supplemental Digital Content 3, http://links.lww.com/RLI/A236, Fig. 7 U), which is comparable to the simulations [27] and agrees with the observation that at least 2 b-values in the domain $0 < b < 50$ mm$^{-2}s$ are required to determine an unbiased $D^*$ [44]. Dyvorne et al. [45] showed that in their data, when $D^*$ was fitted only 4 b-values were needed for IVIM in the liver. However, their approach is different as they did not take into account the increase in wCV when fewer b-values are used. We favor this criterion since in longitudinal studies that aim to detect treatment effects or disease progression, the reproducibility directly relates to the statistical power.

In the literature, it is common to measure b-values up to $b = 800$ to $1100$ mm$^{-2}s$ [5, 9, 13, 14, 17, 22, 23, 25, 33]. In addition, simulations show that high b-values help
to determine $D$ [27]. However, including higher $b$-values to the protocol increases the $TE$ of the acquisitions of all $b$-values, and, consequently, leads to SNR decreases for all DW images. This effect was not taken along in previous simulations and cannot be undone by deleting data after the acquisition. Therefore, we chose to limit our maximum $b$-value to 500 mm$^{-2}$s. As our $D$ value was similar to values reported in the literature, we believe limiting the highest $b$-values did not introduce systematic errors. Including higher $b$-values would potentially help improve $wCV_D$, but will worsen $wCV_f$ due to the loss of SNR. As our $wCV_D < wCV_f$ and our $wCV_D$ is comparable to $wCV_D$'s in the liver found in the literature (5%–20%) [13, 23, 34, 45, 46], there is no need for $b$-values higher than 500 mm$^{-2}$s for IVIM in these organs.

The intrasession $wCV$s (7%–12% for $D$ and 19%–36% for $f$, Table 7.4) were not significantly lower than the intersession $wCV$s, (5%–13% and 23%–55%). This suggests that long term variations are small. Possibly, the difference between inter and intra $wCV$s becomes even smaller when scanning fasted patients at the same time during the day to minimize the physiological fluctuations.

Typically, for normally distributed acquisition noise, one would expect that reproducibility should increase with increasing scan time. The fact that we could decrease the measurement time substantially at the cost of only a limited loss in intersession $wCV$s, indicates that other factors that do not decrease with additional acquisitions play a role as well. Such factors could include the presence of image artifacts, the mismatch between the IVIM model and the physiology of liver and pancreas, and short term physiological changes. When patients show pathologies, it may be easier to delineate ROIs and ROI placement could become more reproducible. In these cases, longer acquisitions may still improve the $wCV$s more substantially than shown in our data.

In the literature of IVIM, multiple measures for repeatability are used [13, 22, 23, 34, 45]. We observed the following measures throughout the literature: the $wCV$ as in this study, the mean of the coefficient of variations calculated per patient, the interscan error and the 95% confidence limits. In this research, we used the method discussed by Barnhart et al. [42], which is in agreement with Bland and Altman [47]. Ignoring the different ways to calculate $wCV$s, the range of published $wCV$s for IVIM in the liver are $wCV_f = 10\%$ to $32\%$ and $wCV_D = 5\%$ to $20\%$ when large ROIs (4–27 cm$^3$) are used [13, 23, 34, 45, 46]. While $wCV$s in our ROI$_{large}$ were in the same range, the $wCV$s in our ROI$_{small}$ were somewhat larger. We expect this is due to our ROI$_{small}$ (1.7 cm$^3$) being smaller than the typical ROIs used so far.

A limitation of this study is that we did not include data from patients with pancreatic or liver pathologies. $wCV$s in patients may differ from our $wCV$s due to various reasons, for example, patients may have biliary duct stents or other metal
devices that disrupt the DWI signal [48], underlining the need of investigating the reproducibility in a separate patient cohort. In addition, $D^*$ may be different in pathologies than in healthy tissue, and the effect of fixing $D^*$ to a reference value from healthy volunteers should be investigated in this case. Furthermore, the bias was deduced using surrogate reference values obtained from the full acquisition. Potentially, there are systematic errors in those reference values and, therefore, the true bias may be higher. However, as $|\text{ND}|$ remains low for the acquisition schemes suggested, we can conclude that the increase in bias from the full data set to our optimized set, is small, and thus scanning longer will not decrease bias. Finally, our protocol only covered 84 mm in cranial-caudal direction, as the protocol was initially implemented for pancreas imaging. The liver is often larger in the cranial-caudal direction and increasing the number of slices, or the slice thickness is needed when the evaluation of the entire liver is desired.

In the literature, there are many different fitting algorithms used for the IVIM model, and the introduced reproducibility and bias in fit parameters will depend on the fitting algorithm [49]. Barbieri et al. [46] showed that a Bayesian probability based algorithm improves the reproducibility and decreases the bias of their IVIM fits. However, the behavior of fitting algorithms may also depend on the number of b-values and acquisitions per b-value used. This falls outside the scope of this research and we restricted ourselves to the Levenberg-Marquardt least squares fit, which is widely available.

To select the ideal settings for reproducible and unbiased IVIM modeling we selected an acquisition that was 20% less reproducible than the full acquisition. Other users may prefer other criteria that can lead to other ideal settings. With the use of the heat maps (see Supplementary Graphs, Supplemental Digital Content 2, http://links.lww.com/RLI/A235 and Supplemental Digital Content 3, http://links.lww.com/RLI/A236), ideal settings for each specific purpose can be deduced.

The bCVs found were similar to the wCVs of the full dataset. This suggests that typical differences between healthy volunteers are lower than the reproducibility of the acquisition protocol. Therefore, for pathologies to be detectable in single subjects with IVIM, the changes need to be larger than the bCVs and wCVs found in this research, which, for some pathologies, is indeed the case. For instance, when the IVIM model is used for lesion classification in the pancreas, $f$ is 72% to 97% lower in pancreas carcinoma compared with healthy pancreas [5, 9], except in the case of neuroendocrine tumors, in which the IVIM model was not conclusive [9, 17]. In hepatocellular carcinoma, a decrease in $f$ of 64% on average was reported when compared with healthy liver tissue [8]. Besides, it was shown that $f$ changes by 79% between hypo-vascular and hyper-vascular hepatic focal lesions [10]. Moreover,
when compared with healthy liver tissue, $D$ is higher by 5% to 29% in hepatocellular carcinoma, 10% in colorectal metastasis, 53% in hemangioma, and 120% in cysts [8, 22, 23]. These values in majority lay above the $bCV$s and $wCV$s from our optimized acquisitions, so these pathologies may be detectable in individual patients using our IVIM protocol. It has also been shown that IVIM-like fits with only three $b$-values can offer diagnostic value in pancreas cancer patients [50]. Finally, in fibrotic liver, $f$ decreases 23% to 40% when compared with healthy livers, whereas $D$ decreases 7% to 18% [13, 14, 25]. These values are in the order of the $bCV$s and $wCV$s from our full acquisition for ROI_large. Therefore, even though the IVIM model shows a difference between the healthy volunteer group and the patient group, it will be challenging to classify liver fibrosis for individual patients.

The scan times found in this research are shorter than scan times of typical clinical respiratory compensated scan protocols (mean: 8:30 minutes) [5, 6, 11, 13, 14, 22, 25, 26]. In addition, scanning with the suggested $b$-values should perform better in terms of reproducibility than other protocols with identical acquisition time, but different $b$-values. Finally, our $wCV$s are lower than typical clinically relevant changes. Therefore, we believe that the protocol as presented here, with a known magnitude of bias and reproducibility, should help clinicians introduce DWI for IVIM modeling in the clinic.

**Conclusion**

Intravoxel incoherent motion model data can be obtained using 7 $b$-values and 6 acquisitions per $b$-value when the entire organ is of interest, and with 7 $b$-values and 9 acquisitions per $b$-value when measuring in a lesion. For the liver, IVIM can be obtained using 3 $b$-values and 9 acquisitions per $b$-value. Using these settings, reproducibility decreased less than 20% from the full acquisition of 14 $b$-values and 9 acquisitions per $b$-value. The understanding of the behavior of reproducibility and bias as a function of $b$-values and acquisitions per $b$-values scanned will help researchers select the shortest IVIM protocol.
References


Supplemental Digital Content 1

wCVs
The wCVs were calculated as defined by Barnhart et al. (2009), but as percentage: the within subject standard deviation divided by the mean parameter value multiplied by 100%. The mean parameter value is averaged both over subjects and measurements. The within subject standard deviation was calculated as the root of the within subject mean of squares. The within subject mean of squares, in our situation, was the mean (over the volunteers) of the squared difference of the repeated measure divided by two.

Supplemental Digital Content 2

Biexponential model, fixed D*
Intrasession
The mean parameter values for intrasession (Table 7.A), and thus used as reference values for ND, where similar to those for intersession (Table 7.2).

Table 7.A. Average value found for D, f and D* over all volunteers using the full data set from the intrasession scans with D*-fixed and D*-free.

<table>
<thead>
<tr>
<th></th>
<th>D*-fixed</th>
<th>D*-free</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>D±SD (10⁻³ mm²s⁻¹)</td>
<td>f±SD (%)</td>
</tr>
<tr>
<td>Pancreas</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ROI_{large}</td>
<td>1.40±0.17</td>
<td>5.1±1.2</td>
</tr>
<tr>
<td>ROI_{small}</td>
<td>1.32±0.16</td>
<td>5.9±3.0</td>
</tr>
<tr>
<td>Liver</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ROI_{large}</td>
<td>1.12±0.09</td>
<td>4.7±1.6</td>
</tr>
<tr>
<td>ROI_{small}</td>
<td>0.10±0.14</td>
<td>5.0±1.4</td>
</tr>
</tbody>
</table>
CHAPTER 7

Figure 7.A. Heat maps of the intrasession $wCV_D$ plotted as function of the number of acquisitions per b-value (x-axis) and number of b-values (y-axis) taken along. Heat maps are shown for the pancreas (left) and the liver (right) for several under sampling schemes per organ (low, high and random). Also, the $wCV$ is plotted for ROI$_{large}$ (top row), and ROI$_{small}$ (bottom row). White voxels indicate that no value was obtained for this combination of b-values and averages. Color version is available at http://links.lww.com/RLI/A235.

Figure 7.B. Heat maps of the intrasession $wCV_f$ plotted as function of the number of acquisitions per b-value (x-axis) and number of b-values (y-axis) taken along. The grouping in this figure is similar to Fig. 7.A. Color version is available at http://links.lww.com/RLI/A235.
Figure 7.C. Heat maps of the intrasession $ND_D$ plotted as function of the number of acquisitions per b-value (x-axis) and number of b-values (y-axis) taken along. The grouping in this figure is similar to Fig. 7.A. Color version is available at http://links.lww.com/RLI/A235.

Figure 7.D. Heat maps of the intrasession $ND_I$ plotted as function of the number of acquisitions per b-value (x-axis) and number of b-values (y-axis) taken along. The grouping in this figure is similar to Fig. 7.A. Color version is available at http://links.lww.com/RLI/A235.
Intersession
The intersession graphs for wCVs and NDs (Fig. 7.E-H) are similar to the intrasession graphs. The equivalent of Table 7.A for intersession is found as Table 7.1 in the main manuscript.

Figure 7.E. Heat maps of the Intersession wCV₀ plotted as function of the number of acquisitions per b-value (x-axis) and number of b-values (y-axis) taken along. The grouping in this figure is similar to Fig. 7.A. Color version is available at http://links.lww.com/RLI/A235.

Figure 7.F. Heat maps of the Intersession wCV₁ plotted as function of the number of acquisitions per b-value (x-axis) and number of b-values (y-axis) taken along. The grouping in this figure is similar to Fig. 7.A. Color version is available at http://links.lww.com/RLI/A235.
Figure 7.G. Heat maps of the Intersession $ND_D$ plotted as function of the number of acquisitions per b-value (x-axis) and number of b-values (y-axis) taken along. The grouping in this figure is similar to Fig. 7.A. Color version is available at http://links.lww.com/RLI/A235.

Figure 7.H. Heat maps of the Intersession $ND_f$ plotted as function of the number of acquisitions per b-value (x-axis) and number of b-values (y-axis) taken along. The grouping in this figure is similar to Fig. 7.A. Color version is available at http://links.lww.com/RLI/A235.
Supplemental Digital Content 3

Biexponential model, free D*

Intrasession

When $D^*$ was set as fitting parameter, the fitted value for $D$ and $f$, averaged over all volunteers, where similar to the fixed $D^*$ model when all b-values and measurements were taken along (Table 7.A-B). However, $wCV$s and $ND$s became worse, especially when data is deleted. Therefore, when precise IVIM model fitting with high reproducibility is desired with $D^*$ as fitting variable, longer acquisition times are required, in which more b-values and measurements per average are acquired than when $D^*$ is fixed.

Table 7.B. Inter- and intrasession coefficients of variation using 14 b-values and 9 measurements per b-value.

<table>
<thead>
<tr>
<th></th>
<th>Intra</th>
<th>Inter</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$wCV_D$ (%)</td>
<td>$wCV_f$ (%)</td>
</tr>
<tr>
<td>Pancreas</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ROI_{large}</td>
<td>8</td>
<td>33</td>
</tr>
<tr>
<td>ROI_{small}</td>
<td>14</td>
<td>40</td>
</tr>
<tr>
<td>Liver</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ROI_{large}</td>
<td>10</td>
<td>45</td>
</tr>
<tr>
<td>ROI_{small}</td>
<td>16</td>
<td>51</td>
</tr>
</tbody>
</table>

Figure 7.I. Heat maps of the intrasession $wCV_D$ plotted as function of the number of acquisitions per b-value (x-axis) and number of b-values (y-axis) taken along. $D^*$ was set as a variable instead of a constant in the fits. The grouping in this figure is similar to Fig. 7.A. Color version is available at http://links.lww.com/RLI/A236.
Figure 7.J. Heat maps of the intrasession wCV, plotted as function of the number of acquisitions per b-value (x-axis) and number of b-values (y-axis) taken along. D* was set as a variable instead of a constant in the fits. The grouping in this figure is similar to Fig. 7.A. Color version is available at http://links.lww.com/RLI/A236.

Figure 7.K. Heat maps of the intrasession wCV, plotted as function of the number of acquisitions per b-value (x-axis) and number of b-values (y-axis) taken along. The grouping in this figure is similar to Fig. 7.A. Color version is available at http://links.lww.com/RLI/A236.
Figure 7.L. Heat maps of the intrasession ND, plotted as function of the number of acquisitions per b-value (x-axis) and number of b-values (y-axis) taken along. D* was set as a variable instead of a constant in the fits. The grouping in this figure is similar to Fig. 7.A. Note the change in scale when compared to Figs C and G. Color version is available at http://links.lww.com/RLI/A236.

Figure 7.M. Heat maps of the intrasession ND, plotted as function of the number of acquisitions per b-value (x-axis) and number of b-values (y-axis) taken along. D* was set as a variable instead of a constant in the fits. The grouping in this figure is similar to Fig. 7.A. Note the change in scale when compared to Figs D and H. Color version is available at http://links.lww.com/RLI/A236.
Interpersonal

Intersession wCVs and NDs show similar behavior as the intrasession versions (Fig. 7.O-T). When 20% larger wCVs are allowed, precise IVIM model fits with high reproducibility for ROI\textsubscript{large} can be done in 8 minutes for the pancreas using 12 b-values and 8 acquisitions per b-value and 5:30 minutes for the liver using 9 b-values and 6 acquisitions per b-value. For ROI\textsubscript{small}, it can be done in 8 minutes for the pancreas using 12 b-values and 8 acquisitions per b-value and 5:15 minutes for the liver using 9 b-values and 7 acquisitions per b-value (Fig. 7.U).

Figure 7.O. Heat maps of the Intersession wCV\textsubscript{D*} plotted as function of the number of acquisitions per b-value (x-axis) and number of b-values (y-axis) taken along. D\textsuperscript{*} was set as a variable instead of a constant in the fits. The grouping in this figure is similar to Fig. 7.A. Color version is available at http://links.lww.com/RLI/A236.
Figure 7.P. Heat maps of the Intersession wCV$_f$ plotted as function of the number of acquisitions per b-value (x-axis) and number of b-values (y-axis) taken along. $D^*$ was set as a variable instead of a constant in the fits. The grouping in this figure is similar to Fig. 7.A. Color version is available at http://links.lww.com/RLI/A236.

Figure 7.Q. Heat maps of the Intersession wCV$_{br}$ plotted as function of the number of acquisitions per b-value (x-axis) and number of b-values (y-axis) taken along. The grouping in this figure is similar to Fig. 7.A. Color version is available at http://links.lww.com/RLI/A236.
Figure 7.R. Heat maps of the Intersession ND\(_D\) plotted as function of the number of acquisitions per b-value (x-axis) and number of b-values (y-axis) taken along. D\(_D^*\) was set as a variable instead of a constant in the fits. The grouping in this figure is similar to Fig. 7.A. Note the change in scale when compared to Figs 7.C and 7.G. Color version is available at http://links.lww.com/RLI/A236.

Figure 7.S. Heat maps of the Intersession ND\(_l\) plotted as function of the number of acquisitions per b-value (x-axis) and number of b-values (y-axis) taken along. D\(_l^*\) was set as a variable instead of a constant in the fits. The grouping in this figure is similar to Fig. 7.A. Note the change in scale when compared to Figs 7.D and 7.H. Color version is available at http://links.lww.com/RLI/A236.
Figure 7.T. Heat maps of the Intersession ND\textsubscript{p}, plotted as function of the number of acquisitions per b-value (x-axis) and number of b-values (y-axis) taken along. The grouping in this figure is similar to Fig. 7.A. Color version is available at http://links.lww.com/RLI/A236.

Figure 7.U. Plots of the scan times as function of additional relative error on intersession wCV using the ROI\textsubscript{large} (left) and ROI\textsubscript{small} (right).
Supplemental Digital Content 4

Simulations
To study the behavior of the fit parameters as a function SNR, we used simulations in Mathematica (Wolfram Research, Champaign, IL). We simulated typical IVIM-data using all b-values from this research and the parameters found from the fit to all volunteer data (pancreas, ROI\textsubscript{large}). Gaussian noise was added to the data for several SNR values (range 2–100). Then, we fitted the IVIM model to the data. This procedure was repeated 1000 times and the mean IVIM model parameters per SNR level were calculated and plotted as a function of SNR (Fig. 7.V). It is clear that systematic errors are introduced at low SNR.

Figure 7.V. IVIM model parameter values from fits on simulated data, as a function of SNR. The stripe indicates the reference value used in the simulation. Note that $D^*$ is plotted on a logarithmic scale.