Journal of Associated Medical Sciences 2027; 60(1): 268-281

JAMS
Journal of Associated Medical Sciences

Scopus Indexed & Thai-Journal Citation Index Centre (TCI)

Journal of Associated Medical Sciences

Journal homepage: https://www.tci-thaijo.org/index.php/bulletinAMS/index



Optimization of IVIM-Diffusion MRI protocol for quantifying pancreatic perfusion between free breathing and respiratory triggering controls

Phimphitcha Ratiphunpong*, Kanyanat Sukiam, Chompoonut Arsakij, Nicharat Klaewsarikorn, Monchai Phonlakrai

School of Radiological Technology, Faculty of Health Science Technology, Chulabhorn Royal Academy, Bangkok, Thailand.

ARTICLE INFO

ORCID of all authors

Phimphitcha Ratiphunpong: https://orcid.org/0009-0006-1551-3417
Kanyanat Sukiam: None
Chompoonut Arsakij: None
Nicharat Klaewsarikorn: None
Monchai Phonlakrai: https://orcid.org/0000-0002-7335-426X

Correspondence

Phimphitcha Ratiphunpong
School of Radiological Technology, Faculty of Health Science Technology, Chulabhorn Royal Academy, Bangkok, Thailand.
Email: Phimphitcha.rat@cra.ac.th

Article history

Received 6 June 2026
Accepted as revised 26 October 2026
Available online 12 November 2026

Keywords

IVIM-DWI, ADC, MRI, water diffusion, perfusion, repeatability

doi: 10.66285/JAMS.2027.025
E-ISSN: 2539-6056

ABSTRACT

Background: Pancreatic steatosis and cancer are major health concerns linked to obesity and metabolic disorders. Because these conditions may alter pancreatic tissue microstructure and microcirculation, non-invasive imaging biomarkers capable of characterizing such changes are needed. Intravoxel incoherent motion diffusion-weighted imaging (IVIM-DWI) enables non-contrast quantification of diffusion and perfusion parameters, including the apparent diffusion coefficient (ADC), true diffusion coefficient (D_{slow}), pseudo-diffusion coefficient (D_{fast}), and perfusion fraction (f). However, the repeatability of such parameters remains a limitation in clinical use.

Objectives: This study aims to evaluate and compare the repeatability of Bayesian-based IVIM-DWI for pancreatic perfusion quantification using free-breathing and respiratory-triggered acquisitions at both global and voxel levels.

Materials and methods: The study consisted of two parts: 1) an evaluation using the NIST/QIBA diffusion phantom to compare image quality and ADC variability between protocols using 10 and 13 b-values, and 2) an assessment in twenty young healthy volunteers to compare within-subject coefficient of variation (wCV) of IVIM-DWI parameters of the selected protocol from the first step between breathing control techniques of respiratory triggering (RT) and free-breathing (FB) using a 3.0 Tesla MRI scanner.

Results: The phantom study demonstrated that the protocol with 13 b-values provided superior image quality and significantly lower ADC variability than the one with 10 b-values ($p<0.05$). Meanwhile, in the volunteer study, the wCV of IVIM-DWI parameters did not significantly differ between RT and FB techniques ($p>0.05$). However, RT exhibited lower wCV for D_{slow} and f, indicating better repeatability.

Conclusion: The preliminary results showed that the 13 b-value IVIM-DWI protocol has potential for evaluating pancreatic perfusion and offers a stable estimation of the perfusion-related f parameter. However, further study is required to validate the agreement between this parameter and gadolinium-based contrast-enhanced MRI.

Introduction

Pancreatic steatosis, first described in 1933, refers to the accumulation of fat within pancreatic tissue, primarily due to obesity. This condition adversely affects pancreatic blood flow via the fat deposits directly compressing the blood vessels and inducing inflammation, which subsequently leads to fibrosis. These processes contribute to changes in the pancreatic vasculature, metabolic dysfunctions, and the progressive replacement of pancreatic parenchyma with adipose tissue. These factors ultimately increase the risk of diseases, such as type 2 diabetes mellitus and pancreatitis, that may potentially progress to pancreatic cancer.[1]

The diagnosis of pancreatic diseases relies heavily on medical imaging modalities. Radiological imaging, in particular, plays a crucial role in evaluating early-stage tumors to guide surgical planning and improve patient outcomes.[2] Among the available imaging modalities, techniques capable of assessing pancreatic perfusion are of particular interest because perfusion differences may aid in the detection and differentiation of pancreatic diseases and in monitoring treatment response. Examples of these include transabdominal ultrasound,[3] positron emission tomography (PET),[4] computed tomography (CT),[5] and magnetic resonance imaging (MRI). Especially, dynamic contrast-enhanced MRI (DCE-MRI), often combined with the administration of a contrast agent, remains a widely used technique for detecting small pancreatic tumors and assessing their vascular invasion.[6] However, DCE-MRI is of limited value in patients with impaired renal function who are contraindicated for contrast agents.

To overcome this limitation, intravoxel incoherent motion diffusion-weighted imaging (IVIM-DWI) has been introduced.[7] IVIM-DWI enables functional imaging of the pancreas without the need for contrast agents.[8] This technique simultaneously captures true molecular diffusion and pseudo-diffusion related to microvascular blood flow. Using a biexponential mathematical model, IVIM-DWI separates these components and provides quantitative parameters that are valuable for assessing apparent diffusion coefficient (ADC), pure molecular diffusion coefficient (D_{slow}), pseudo-diffusion coefficient (D_{fast}), and perfusion fraction (f).[9]

Recent studies have demonstrated the utility of IVIM-DWI for characterizing pancreatic lesions. In particular, the f parameter has been shown to differ significantly between pancreatic cancer patients and individuals without pancreatic cancer, highlighting IVIM-DWI's potential as a noninvasive imaging tool for identifying pathological conditions of the pancreas.[10] Nevertheless, the absence of a standardized IVIM-DWI acquisition and analysis protocol among research sites remains a major limitation. Variations in echo time (TE), motion control strategies (e.g., breath-hold, free-breathing, respiratory triggering),[11] MRI system, IVIM-fitting method, and magnetic field strength[12] can lead to inconsistent parameter estimation and poor repeatability. Moreover, respiratory motion remains a significant source of artifacts, adversely impacting the accuracy of IVIM measurements.[13] Few studies have combined b-value optimization with respiratory motion-control comparison for pancreatic IVIM-DWI within a single design.[11,12]

To address these challenges, the present study was established to develop and optimize an IVIM-DWI protocol for pancreatic imaging. Initially, the selection of b-values was refined using a phantom experiment evaluating technical diffusion performance (SNR and ADC variability). The technically optimized b-value scheme was then applied to healthy volunteers, comparing image quality and parameter repeatability between respiratory-triggered (RT) and free-breathing (FB) acquisition techniques using a 3.0 Tesla (T) MRI scanner. This study is an initial stage of investigation to ultimately establish a reliable IVIM-DWI protocol suitable for future clinical research on pancreatic perfusion.

Materials and methods

Study design

This experimental study was planned with the goal of evaluating the repeatability of IVIM-DWI for quantifying pancreatic perfusion in healthy volunteers. The research included both in vitro and in vivo components. The in vitro study employed the NIST/QIBA diffusion phantom to assess image quality and parameter variability using two schemes with different numbers of b-values. The in vivo study investigated repeatability for two respiratory control methods in a cohort of 20 healthy subjects.

In vitro study (phantom)

The in vitro component of this work evaluated the image quality and parameter variability derived from IVIM-DWI using two b-value schemes (10 and 13 b-values) with a 3.0 T MRI scanner. The NIST/QIBA Diffusion Phantom, designed for quantitative diffusion MRI, was used for this purpose. This phantom contains sample tubes filled with polyvinylpyrrolidone (PVP) solutions at concentrations of 10% to 50% alongside reference tubes containing distilled water. This configuration mimics varying degrees of diffusion, like biological tissues. Specifically, the phantom experiment was designed to evaluate the technical performance of two b-value schemes in terms of image quality (SNR) and ADC's coefficient of variation (CV), to inform b-value scheme selection for the subsequent in vivo study. Direct validation of IVIM-specific parameters (D_{slow}, D_{fast}, f) against a phantom reference standard was not performed, as no established IVIM phantom standard was available.

MRI acquisition

IVIM-DWI data from the phantom were acquired using a Philips Ingenia 3.0T MRI system (Philips Healthcare, Best, The Netherlands) with an 18-channel body coil. All scans were conducted at room temperature (16±3 °C). Two separate IVIM-DWI protocols using 10 or 13 b-values were employed. Both b-value schemes were selected based on two considerations: the recommendation from a previous study[14] to use at least seven b-values for reliable and clinically acceptable measurements, and the need to maintain a clinically acceptable scan time.

Each acquisition was repeated three times continuously without repositioning the phantom to assess repeatability. The parameters of the IVIM-DWI protocol are summarized in Table 1.

Table 1. IVIM-DWI scanning protocol for phantom study.

Parameter	10 b-values	13 b-values
Sequence	SE-EPI (DwiSE)	SE-EPI (DwiSE)
Fat suppression	SPIR	SPIR
Parallel imaging [a]	SENSE (factor 2)	SENSE (factor 2)
Slice	30	30
Phase Encoding Dir.	A>>P	A>>P
Phase Oversampling	0%	0%
FOV (mm)	400×400	400×400
Slice thickness (mm)	4	4
Gap (mm)	0	0
Acquisition Voxel size (mm^3)	2.5×2.54×4.0	2.5×2.54×4.0
Reconstructed Voxel size (mm^3)	1.08×1.08×4.0	1.08×1.08×4.0
Acquisition matrix (mm^2)	152×129	152×129
TR (ms)	1,799	1,798
TE (ms)	86	86
Diffusion weightings	10	13

Table 1. IVIM-DWI scanning protocol for phantom study (continued).

Parameter	10 b-values	13 b-values
b-values (NSA)(s/mm^2)	0(1), 10(2), 30(2), 50(2), 70(2), 100(2), 150(2), 200(3), 400(3), 800(3)	0(1), 10(2), 30(2), 40(2), 50(2), 60(2), 70(2), 80(2), 100(2), 150(2), 200(3), 400(3), 800(3)
Distortion correction	2D	2D
Scan time	3.57 mins	5.02 mins

Note: [a]Acceleration factor in phase direction. A>>P: anterior to posterior, EPI: echo-planar imaging, FOV: field of view, NSA: number of signal averages, SE: spin echo, SENSE: sensitivity encoding, SPIR: spectral pre-saturation with inversion recovery, TE: echo time, TR: repetition time, ms: millisecond.

Image processing and quantification

The IVIM-DWI datasets from the two-imaging acquisition protocol were utilized to generate S_0 maps using IVIMmodelfit toolbox implemented in MATLAB R2022b (MathWorks, Natick, MA, USA), as described elsewhere.[15,16] The ADC maps were generated using mono-exponential fitting on an MRI workstation using Philips IntelliSpace Portal software (Version 5.3.1.2, Philips Healthcare, Best, The Netherlands). Subsequently, regions of interest (ROIs) were drawn for each PVP concentration to derive the mean and standard deviation (SD) of the signal-to-noise ratio (SNR) on the S_0 maps; the within-session variability and percent coefficient of variation (%CV) for the ADC value of each PVP concentration were also calculated. The SNR was calculated using the mean signal divided by the noise SD, as described in Equation 1.

$$SNR = \frac{Mean_{signal}}{SD_{noise}} \quad (1)$$

Where $Mean_{signal}$ is the average signal intensity within the ROI and SD_{noise} is the standard deviation of signal in a background

For each PVP concentration, repeatability of the ADC measurement was assessed across the three repeated acquisitions using three established metrics following Bland and Altman[17]: the standard deviation (SD), the percent coefficient of variation (%CV), and the repeatability coefficient (RC). SD is standard deviation of the ADC values across the three repeated acquisitions

$$\%CV = \frac{SD}{Mean} x100 \quad (2)$$

$$RC = 2.77 \times SD \quad (3)$$

Where Mean is the mean of the ADC values across the three repeated acquisitions for that concentration. RC represents the smallest difference between two measurements of the same object that can be regarded as real at the 95% level. Because the six PVP concentrations span a wide range of absolute ADC values, a single value across concentrations could not be obtained by direct averaging; %CV was therefore pooled across concentrations as the root-mean-square of the six concentration-specific values, and SD (and hence RC) was pooled as the square root of the mean of the six concentration-specific variances. Ninety-five percent confidence intervals for the pooled SD, %CV, and RC were derived from the chi-squared distribution.

ADC linearity across PVP concentrations was assessed by simple linear regression of mean ADC values against PVP concentration (%), with R^2 reported as the goodness-of-fit index.

In vivo study

Participants

To avoid age-related pancreatic steatosis, a total of 20 young healthy volunteers (4 males, 16 females; age range: 21-36 years; mean age: 23.1±3.74 years) with no history of pancreatic disease were recruited. The inclusion criteria were as follows: age 20-45 years, no prior history of pancreatic disease and no clinical symptoms suggestive of pancreatic dysfunction, and ability to undergo MRI. The exclusion criteria included motion artifacts, incidental abnormalities, and early withdrawal from the study. Ethical approval was granted by the Human Research Ethics Committee of the research site (Approval Code: EC 122/2567). Written informed consent was obtained from all individual participants prior to the MRI scans.

Magnetic resonance imaging acquisition

MRI was performed on healthy volunteers with the same scanner as used for the phantom study. The 13 b-value scheme (Table 1) was selected for all subsequent in vivo acquisitions, based on its superior ADC repeatability demonstrated in the preceding phantom experiment. Each subject underwent a total of four scans: two each for the respiratory triggering (RT) and free-breathing (FB) techniques.

Image processing and quantification

Images were anonymized and processed using 3D Slicer (Version 5.6.2; The Slicer Community, Boston, MA, USA; www.slicer.org) and MATLAB (MathWorks, Natick, MA, USA). Pancreatic segmentation was performed on S_0 map using 3D Slicer's Segment Editor and saved in NIfTI format. Segmentation was carried out by an experienced radiologic technologist with 10 years of expertise in body MRI. Only slices demonstrating the

entire pancreas were included. The whole pancreas-encompassing the head, body, and tail-was delineated slice by slice.

The IVIM-DWI dataset with 13 b-values was used to generate IVIM parameter maps. Bayesian IVIM parameter estimation was performed using the open-source IVIMmodelfit toolbox implemented in MATLAB R2022b (MathWorks, Natick, MA, USA).[15,16]

Bayesian IVIM fitting formula is shown below:

$$S(b) = S_0 [f \cdot \exp(-b \cdot D_{fast}) + (1-f) \cdot \exp(-b \cdot D_{slow})] + \varepsilon_n \quad (4)$$

where S_0 is the signal at b=0 s/mm^2, f is the perfusion fraction, D_{fast} is the pseudo-diffusion coefficient, and D_{slow} is the true diffusion coefficient. ε_n is noise term, typically assumed to follow a Gaussian distribution with variance σ_S^2

For each subject, repeatability of each parameter was assessed across the two repeated scans using the within-subject standard deviation (wSD), within-subject coefficient of variation (wCV), and repeatability coefficient (RC). For each parameter, the difference between scans was defined as $d_i = Scan_{2_i} - Scan_{1_i}$ for subject i.

$$wSD = \frac{SD(d)}{\sqrt{2}} \quad (5)$$

$$wCV(\%) = \frac{wSD}{\bar{x}} x100 \quad (6)$$

$$RC = 2.77 \times wSD \quad (7)$$

Where SD(d) is the standard deviation of the differences across all subjects, and $\bar{x}$ is the grand mean of all scan values across both sessions and all subjects. Because all 20 subjects yielded ADC and IVIM parameter values within a comparable physiological range, wSD, wCV, and RC were each pooled across subjects directly, using the same formulas applied to the full 20-subject dataset. Ninety-five percent confidence intervals for the pooled wSD, wCV, and RC were derived from the chi-squared distribution.

Mean absolute error

To assess voxel-level differences between scans, the mean absolute error (MAE) was calculated using MATLAB for each IVIM parameter and ADC map. The MAE of the pancreatic image volume in all parametric maps was computed across the two scans.

Statistical analysis

***In vitro* study**

The normality of the data distribution was assessed using the Shapiro-Wilk test. Paired t-tests were applied for normally distributed variables, while Wilcoxon signed-rank tests were used for non-normally distributed variables to compare parameter variability between the 10 and 13 b-value schemes. The per-concentration %CV values were compared between the 10- and 13-b-value protocols using the Wilcoxon signed-rank test.

***In vivo* study**

The IVIM and ADC parameter maps were analyzed for mean, SD, and MAE between repeated scans under the breathing control techniques of RT and FB. Repeatability was quantified using the pooled wSD, wCV, and RC, with 95% confidence intervals derived from the chi-squared distribution. Within-technique Bland-Altman analysis (scan 2 - scan 1) was performed separately for FB (FB2 - FB1) and RT (RT2 - RT1) to quantify systematic bias (mean difference) and precision (limits of agreement = bias±1.96×SD[d]). Inter-technique agreement was additionally assessed by Bland-Altman analysis of the mean FB value versus the mean RT value per subject (mean difference and limits of agreement). Shapiro-Wilk tests were used to evaluate the normality of the data. Paired t-tests or Wilcoxon signed-rank tests were applied as appropriate to compare parameters between the RT and FB techniques.

Results

In vitro study (phantom)

The effect of the number of b-values on image quality was first evaluated using the NIST/QIBA diffusion phantom. Visual inspection of the S_0 maps revealed a higher SNR with the 13-b-value protocol than with the 10-b-value protocol across all PVP concentrations. This improvement was particularly evident at low PVP concentrations, at which the signal intensity was higher (Figure 1). Quantitative analysis confirmed significantly higher SNR values with the 13-b-value protocol (p<0.05), suggesting that this protocol was associated with higher image quality.

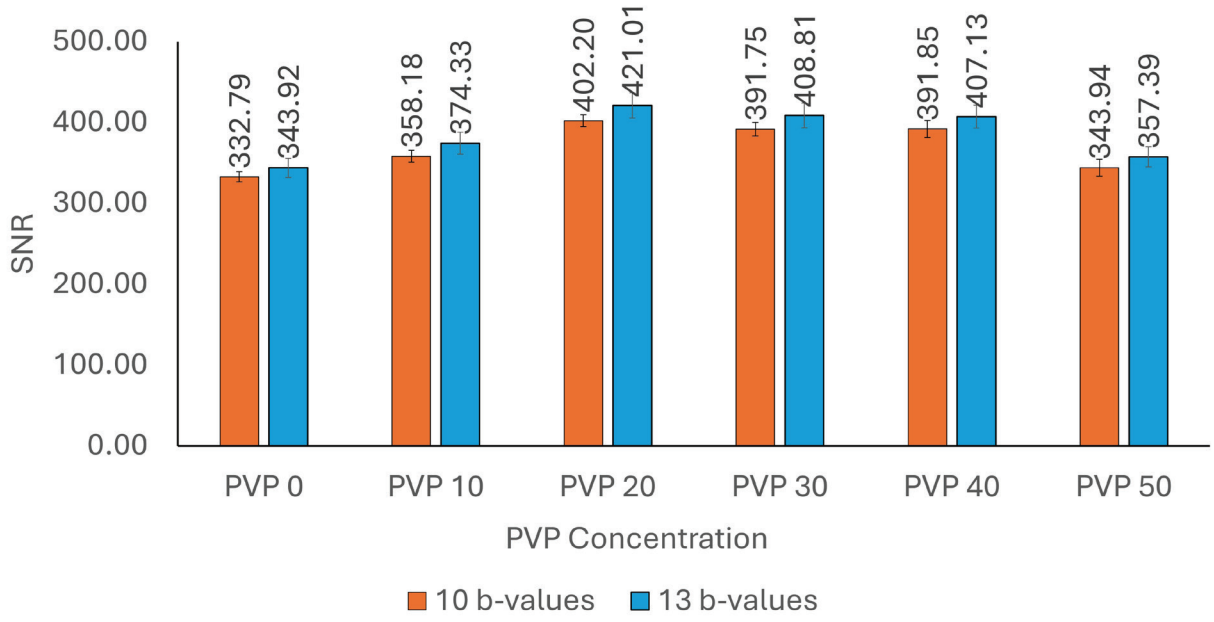


Figure 1. Signal-to-noise ratio values (SNR) derived from S_0 maps, comparing the use of 10 b-values and 13 b-values at each polyvinylpyrrolidone (PVP) concentration level.

Repeatability of ADC measurements was quantified across the three repeated acquisitions, as shown in Table 2. The 13-b-value protocol yielded substantially better precision than the 10-b-value protocol, with pooled %CV of 0.380% (95% CI: 0.272-0.627) versus 1.291% (95% CI: 0.926-2.131), and corresponding repeatability coefficients of 0.009×10^{-3} mm^2/sec (95% CI: 0.006-0.014) versus 0.028×10^{-3} mm^2/sec (95% CI: 0.020-0.047). Per-concentration %CV was lower with 13 b-values at every PVP concentration, and this difference was statistically significant (Wilcoxon signed-rank test, p=0.031). Repeatability coefficients followed the same pattern: RC ranged narrowly from 0.008-0.010×10^{-3} mm^2/sec with 13 b-values, whereas it ranged from 0.016-0.037×10^{-3} mm^2/sec with 10 b-values - more than four-fold higher at low ADC concentrations (PVP 20-30%). This indicates that the smallest detectable real difference in ADC is both smaller and more consistent across the diffusion range when using 13 b-values.

Table 2. Pooled within-session repeatability of phantom ADC measurements for the 10- and 13-b-value protocols (three repeated acquisitions, six polyvinylpyrrolidone (PVP) concentrations). Values in parentheses are 95% confidence intervals.

PVP (%)	13-b-value				10-b-value			
	Mean ADC	SD[a]	%CV[b]	RC[c]	Mean ADC	SD[a]	%CV[b]	RC[c]
0	2.541	0.003	0.13	0.009	2.57	0.006	0.22	0.016
10	2.015	0.003	0.15	0.009	2.04	0.008	0.37	0.021
20	1.591	0.003	0.19	0.008	1.60	0.013	0.83	0.037
30	1.154	0.003	0.30	0.010	1.16	0.013	1.09	0.035
40	0.695	0.003	0.44	0.009	0.69	0.010	1.44	0.028
50	0.416	0.003	0.71	0.008	0.41	0.010	2.42	0.027
Pooled	-	0.0032 (0.002-0.005)	0.38 (0.27-0.63)	0.009 (0.006-0.014)	-	0.0102 (0.0073-0.0168)	1.29 (0.93-2.13)	0.028 (0.020-0.047)

Note: [a]SD across the three repeated acquisitions, ADC and SD in ×10^{-3} mm^2/sec, [b]%CV=SD/mean×100%, Pooled %CV: root mean square of the six per-concentration values, [c]RC: repeatability coefficient = 2.77×SD, in ×10^{-3} mm^2/sec.

For ADC estimation, the 13-b-value protocol consistently yielded lower variability across all PVP concentrations. The within-session %CV for the ADC values across all phantom scanning sessions was notably lower with 13 b-values, indicating better parameter stability (Figure 2).

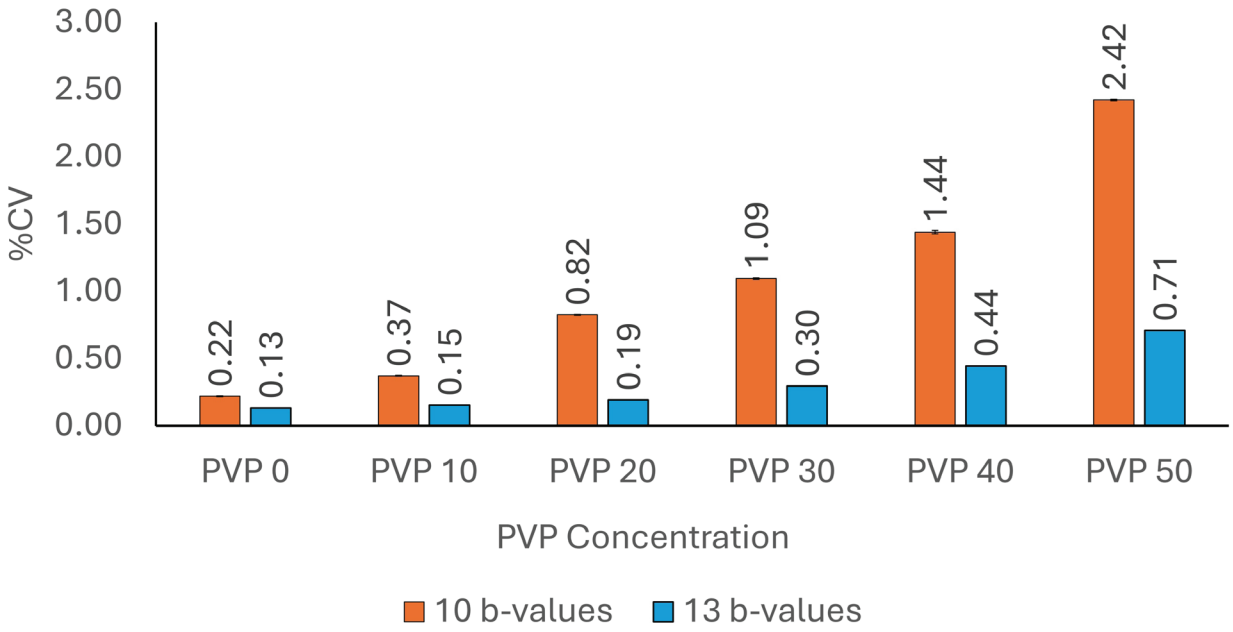


Figure 2. Coefficient of variation (%CV) values calculated from the ADC map derived from the NIST/QIBA diffusion phantom, comparing the use of 10 b-values and 13 b-values at each polyvinylpyrrolidone (PVP) concentration level.

Both protocols demonstrated excellent ADC linearity across PVP concentrations (13b: R^2=0.994, 10b: R^2=0.994, both p<0.001), confirming consistent diffusion sensitivity across the full concentration range.

In vivo study (healthy volunteers)

IVIM-DWI and ADC data were acquired from 20 healthy volunteers to assess the repeatability of pancreatic diffusion parameters, including D_{slow}, D_{fast}, and f, under the RT and FB breathing control techniques.

Repeatability analysis using pooled wSD, wCV, and RC

Pooled repeatability statistics were computed across all 20 subjects (2 scans each) for both acquisition techniques. For D_{slow}, RT demonstrated lower variability than FB: wCV was 6.02% (95% CI: 4.58-8.79%) under RT and 8.38% (95% CI: 6.37-12.24%) under FB, with corresponding RC of 0.224×10^{-3} mm^2/sec (95% CI: 0.170-0.327) and 0.284×10^{-3} mm^2/sec (95% CI: 0.216-0.415). D_{fast} showed the highest variability of all parameters under both techniques: wCV was 15.78% (95% CI: 12.00-23.05%) for FB and 17.48% (95% CI: 13.29-25.53%) for RT, with RC of 17.99×10^{-3} mm^2/sec (95% CI: 13.68-26.28) and 22.63×10^{-3} mm^2/sec (95% CI: 17.21-33.05), respectively. For f, wCV was 11.16% (95% CI: 8.49-16.30%) under FB and 10.20% (95% CI: 7.76-14.91%) under RT, with RC of 9.09% (95% CI: 6.91-13.27) and 8.06% (95% CI: 6.13-11.78). ADC showed the lowest variability: wCV was 4.83% (95% CI: 3.68-7.06%) for FB and 6.04% (95% CI: 4.59-8.82%) for RT, with RC of 0.154×10^{-3} mm^2/sec (95% CI: 0.117-0.225) and 0.202×10^{-3} mm^2/sec (95% CI: 0.153-0.294). Pooled wSD values for all parameters and techniques are provided in Table 3. Boxplots illustrating the distribution of wCV values for each parameter under RT and FB are shown in Figure 3.

Table 3. Parameter values and pooled within-subject repeatability statistics for IVIM-DWI parameters under free-breathing (FB) and respiratory triggering (RT) acquisitions (N=20 subjects, 2 scans per technique).

Parameter	Technique	Mean±SD	wSD [a]	wCV (%) [b]	RC [c]
D_{slow} ($\times10^{-3}$ mm^2/sec)	FB	1.22±0.13	0.102 (0.078-0.150)	8.38 (6.37-12.24)	0.284 (0.216-0.415)
	RT	1.34±0.14	0.081(0.061-0.118)	6.02 (4.58-8.79)	0.224 (0.170-0.327)
D_{fast} ($\times10^{-3}$ mm^2/sec)	FB	41.12±10.47	6.490 (4.936-9.479)	15.78 (12.00-23.05)	17.99 (13.68-26.28)
	RT	46.71±14.80	8.164 (6.209-11.925)	17.48 (13.29-25.53)	22.63 (17.21-33.05)
f (%)	FB	29.38±4.97	3.279 (2.493-4.789)	11.16 (8.49-16.30)	9.088 (6.911-13.274)
	RT	28.51±4.00	2.909 (2.212-4.249)	10.20 (7.76-14.91)	8.064 (6.133-11.778)
ADC ($\times10^{-3}$ mm^2/sec)	FB	1.15±0.09	0.056 (0.042-0.081)	4.83 (3.68-7.06)	0.154 (0.117-0.225)
	RT	1.20±0.13	0.073 (0.055-0.106)	6.04 (4.59-8.82)	0.202 (0.153-0.294)

Note: Mean±SD: mean and standard deviation of per-subject mean values (average of two scans) across all 20 subjects. [a]wSD: pooled within-subject standard deviation = SD ($scan_2$ - $scan_1$) / $\sqrt{2}$, [b]wCV: within-subject coefficient of variation = wSD / grand mean × 100%, Grand mean: mean of all scan values across both sessions, [c]RC: repeatability coefficient = 1.96 × $\sqrt{2}$ × wSD ≈ 2.77 × wSD, defines the smallest real difference at the 95% level. 95% CIs derived from chi-squared distribution. D_{slow}, D_{fast} and ADC in $\times10^{-3}$ mm^2/sec, f in %. Values in parentheses are 95% confidence intervals.

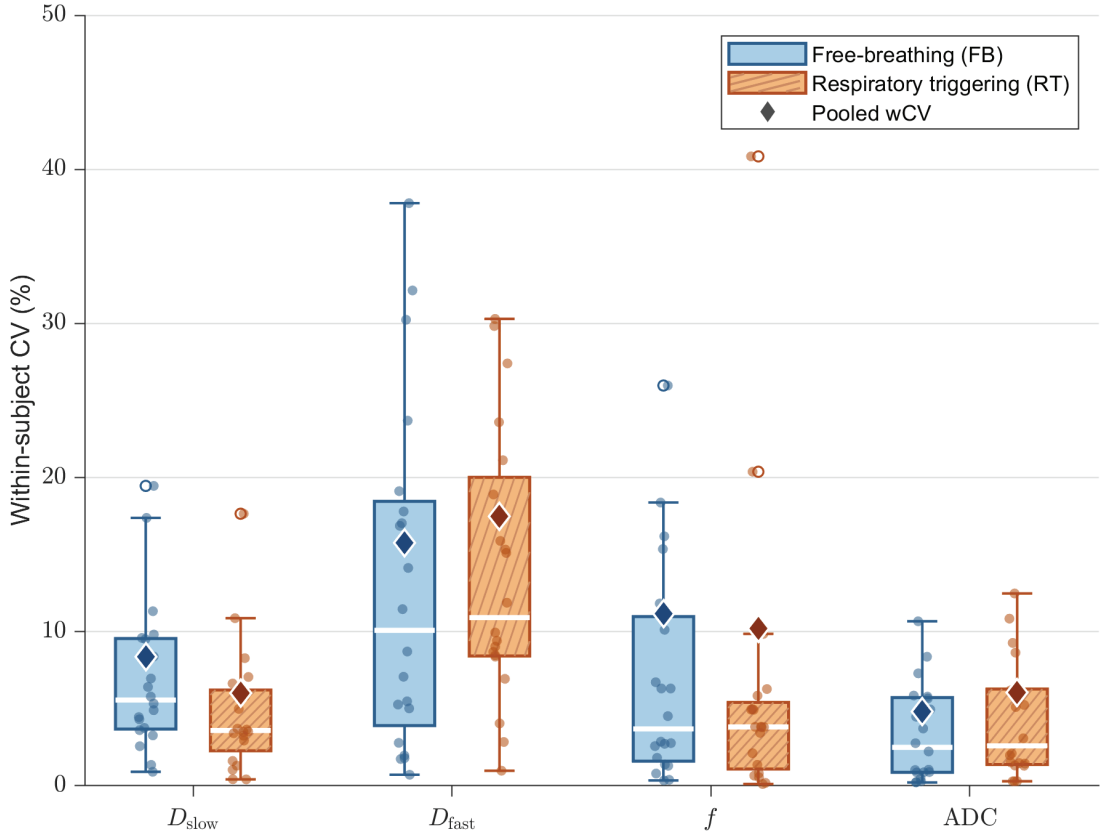


Figure 3. Within-subject coefficient of variation (wCV) of IVIM-DWI parameters under free-breathing (FB) and respiratory triggering (RT) acquisitions. Box plots display the interquartile range (IQR), horizontal white lines indicate the median, and whiskers extend to 1.5×IQR; circles represent individual subject values (N=20). Filled diamonds (♦) indicate the pooled within-subject coefficient of variation (wCV).

Within-technique Bland-Altman analysis

Within-technique Bland-Altman analysis revealed near-zero bias for D_{slow}, f, and ADC under both FB and RT, indicating no systematic drift between repeated acquisitions. D_{fast} showed a modest negative bias under FB (-4.48×10^{-3} mm^2/sec) that was attenuated under RT (-1.11 × 10^{-3} mm^2/sec). Limits of agreement were widest for D_{fast} regardless of technique (FB: [-22.47, +13.51]; RT: [-23.74, +21.52]×10^{-3} mm^2/sec), consistent with its higher wCV. RT yielded narrower limits of agreement (LoA) and lower RC than FB for D_{slow} (RC: 0.224 vs 0.284×10^{-3} mm^2/sec) and f (8.06 vs 9.09%), supporting superior repeatability of triggered acquisition for these parameters. Full Bland-Altman plots are shown in Figure 4.

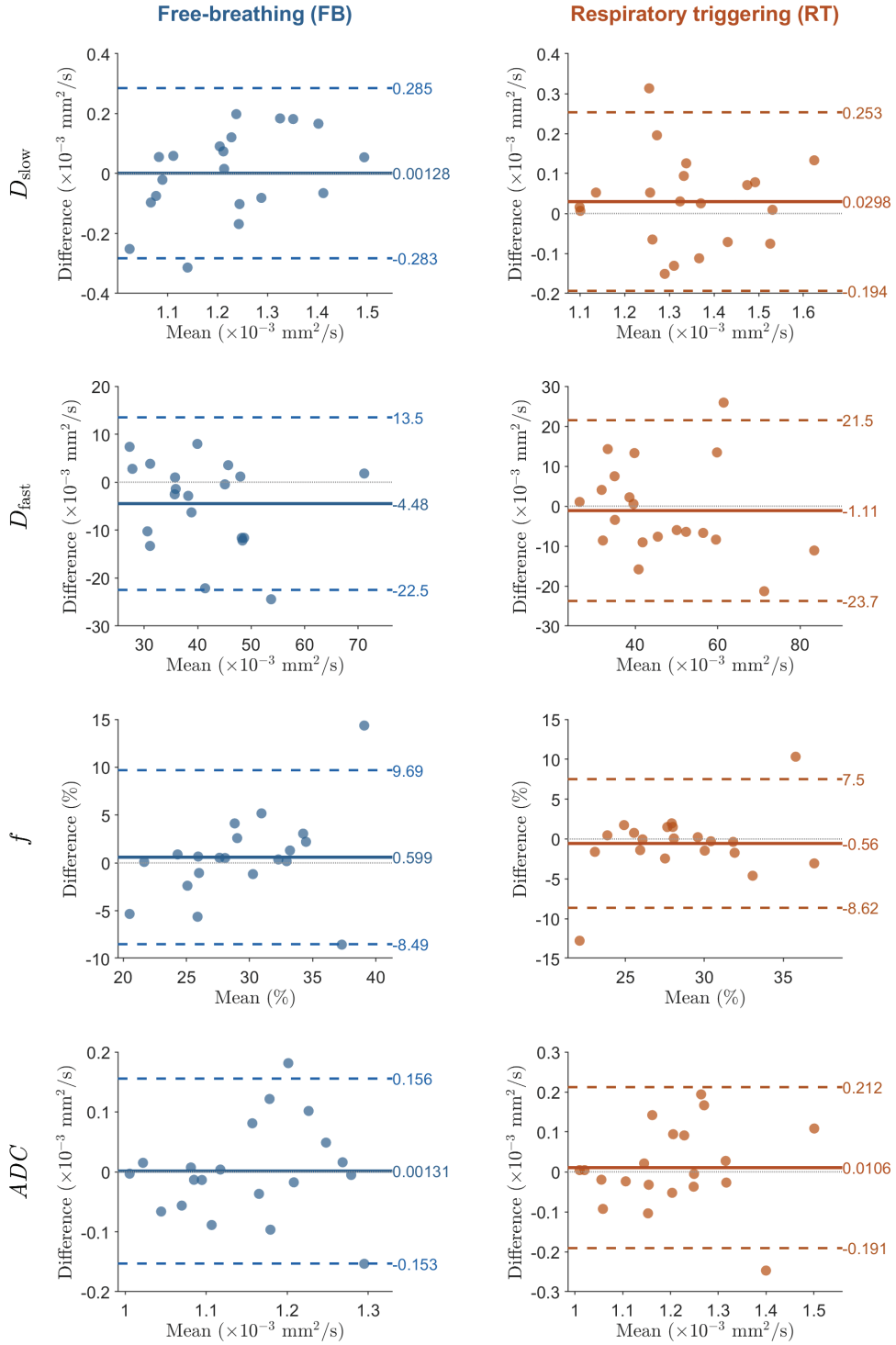


Figure 4. Within-technique Bland-Altman plots for IVIM-DWI parameters under free-breathing (FB, blue) and respiratory triggering (RT, orange). Each panel shows the difference between repeated acquisitions (Scan 2-Scan 1) plotted against their mean for D_{slow}, D_{fast}, f, and ADC. Solid lines indicate mean bias; dashed lines indicate 95% limits of agreement (bias±1.96 SD). Each technique is represented by N=20 subjects.

Agreement between RT and FB: Bland-Altman analysis

Agreement between the RT and FB protocols was evaluated using Bland–Altman plots for each parameter (D_{slow}, D_{fast}, f, and ADC). Mean differences (biases) between FB and RT were -0.12×10^{-3} mm^2/sec for D_{slow}, -5.59×10^{-3} mm^2/sec for D_{fast}, 0.87% for f, and -0.05×10^{-3} mm^2/sec for ADC. The 95% confidence intervals indicated good agreement, with over 90% of the data points falling within the limits of agreement across all parameters (Figure 5). Paired testing of the per-subject means, however, showed that the FB–RT difference was statistically significant for D_{slow} (p<0.001), D_{fast} (p=0.008), and ADC (p=0.014), but not for f (p=0.218), indicating a systematic offset between the two techniques.

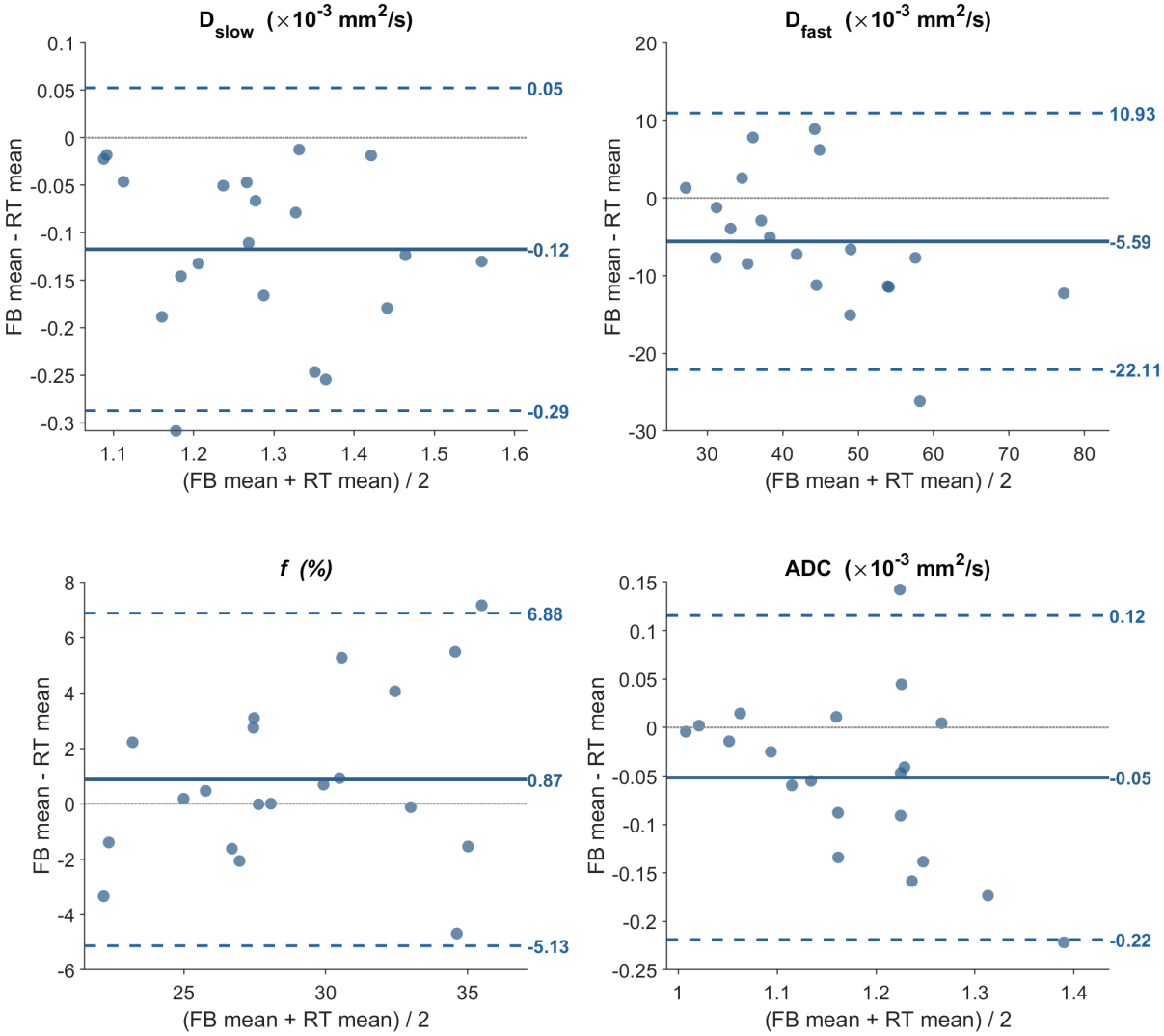


Figure 5. Bland-Altman plots for D_{slow}, D_{fast}, f, and ADC, comparing respiratory triggering (RT) and free-breathing (FB). Each plot displays the mean difference and limits of agreement. Each data point represents one subject (N=20). Solid lines indicate mean bias; dashed lines indicate 95% limits of agreement (bias±1.96 SD).

Statistical comparison of wCV

All four IVIM parameters showed non-normal distributions under at least one technique (p<0.05); therefore, the Wilcoxon signed-rank test was applied throughout. No statistically significant difference in wCV between RT and FB was found for any parameter: D_{slow} (p=0.058), D_{fast} (p=0.756), f (p=0.546), or ADC (p=0.368). Moreover, there was no statistically significant difference in repeatability observed between the RT and FB techniques.

Voxel-based repeatability analysis using MAE

To further evaluate voxel-wise repeatability, the MAE between the first and second scans was computed for each parameter, the results of which are summarized in Table 4. The Shapiro–Wilk test confirmed normality for all MAE distributions (p>0.05). Meanwhile, the paired t-test revealed a statistically significant difference in the MAE of the ADC parameter between RT and FB (p=0.048), while the MAE of D_{slow}, D_{fast}, and f did not show any significant differences. To facilitate visual comparison between RT and FB, the voxel-wise MAE maps for both motion correction techniques for a single subject are presented in Figure 6.

Table 4. Mean and standard deviation (SD) of the mean absolute error (MAE) for IVIM-DWI and ADC parameters obtained from the first and second scans under respiratory triggering and free-breathing techniques.

Parameters	Respiratory triggering MAE±SD	Free breathing MAE±SD	Mean Diff	95% CI	p value
ADC ($\times10^{-3}$ mm^2/sec)	0.22±0.155	0.19±0.131	+0.036	0.000, 0.071	0.048*
D_{slow} ($\times10^{-3}$ mm^2/sec)	0.28±0.063	0.27±0.063	+0.015	-0.008, 0.038	0.199
D_{fast} ($\times10^{-3}$ mm^2/sec)	39.21±11.876	37.63±11.184	+1.581	-2.561, 5.724	0.434
f (%)	11.60±2.606	12.40±3.243	-0.796	-2.035, 0.442	0.194

Note: ADC: apparent diffusion coefficient, D_{slow}: pure diffusion coefficient, D_{fast}: pseudo-diffusion coefficient, f: perfusion fraction, MAE: mean absolute error. *significantly different at p<0.05 using a paired t-test.

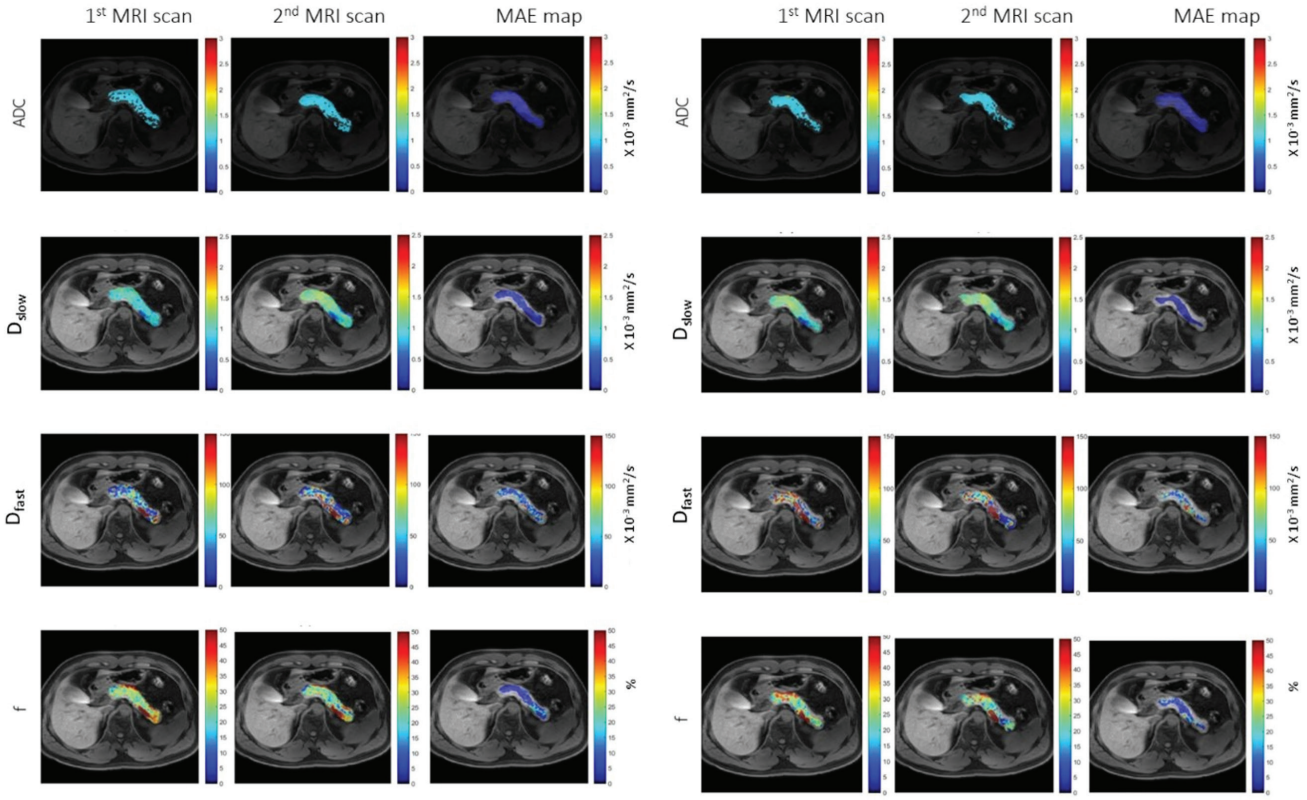


Figure 6. Comparison of voxel-wise MAE maps of ADC, D_{slow}, D_{fast}, and f parameters between scan-rescan acquisitions under respiratory triggering (left panel) and free-breathing (right panel). Each row presents the parameter maps from the first MR scan, the second MR scan, and the corresponding MAE map.

Discussion

The objective of this study was to develop and optimize an IVIM-DWI protocol for measuring pancreatic perfusion, focusing on the selection of appropriate b-values and the assessment of the repeatability of parameters derived from IVIM-DWI and ADC in healthy volunteers. The results showed that in terms of both SNR and %CV, using 13 b-values produced significantly better outcomes than using 10 b-values. These findings support the research hypothesis, which posited that a well-optimized b-value configuration in IVIM-DWI would enhance image quality and ensure high repeatability of ADC measurements at a level sufficient for use in healthy volunteers.

This study revealed that SNR was significantly higher, and %CV was significantly lower, with the use of 13 b-values compared with 10 b-values. SNR was derived from the S_0 map, which is independent of the number of b-values. The additional b-values in the 13-b-value protocol increase diffusion measurements for model fitting rather than S_0 signal averaging. Therefore, the protocols were compared with matched NSA to isolate the effects of the b-value scheme on image quality and parameter precision, rather than total acquisition time.

The pooled RC further supported measurement reliability. The 13-b-value protocol showed consistently narrow RCs ($0.008\text{-}0.010\times10^{-3}$ mm^2/sec), whereas the 10-b-value protocol showed greater variability, particularly at lower ADCs (up to 0.037×10^{-3} mm^2/sec), indicating superior precision with 13 b-values, especially at low diffusion coefficients.

In an *in vivo* study, the results found that 13-b-values improve repeatability (wCV) in normal subjects, which may be beneficial for patients with pancreatic steatosis who require highly accurate analysis.[18,19] The results are also consistent with the study by Gurney-Champion *et al.*,[14] who found that at least 7 b-values IVIM-DWI protocol provided acceptable reproducibility for pancreatic and liver perfusion measurement, with only a 20% reduction compared with a full 14 b-value acquisition.

To compare repeatability between RT and FB, pooled wCV was lower under RT for D_{slow} and f, with corresponding RC values. This could be attributable to the improved respiratory control associated with RT, which minimizes motion artifacts, thereby stabilizing these parameters. In contrast, pooled wCVs of D_{fast} and ADC were higher in RT than in FB, suggesting that RT may not eliminate motion artifacts. This could be due to irregular breathing patterns in some volunteers, which might prolong scan time and increase noise levels; however, noise levels were not directly measured in this study, and this interpretation remains speculative. The D_{fast} parameter also tends to vary more than the others, potentially due to its sensitivity to low SNR and noise present in the b-value range used for its calculation. This leads to greater variability in D_{fast} than in D_{slow} and f. These findings are consistent with those of Liu *et al.*,[20] who found that ADC values demonstrated good consistency across different MRI devices, whereas all IVIM parameters-including D_{fast}, D_{slow}, and f-showed only moderate inter-device reproducibility, with variability attributed primarily to differences in post-processing software and algorithms across devices. Similarly, Nadolska *et al.*[10] reported CV exceeding 50% for D_{fast} across multiple b-value configurations, further confirming that D_{fast} is highly sensitive to acquisition protocol parameters. Meanwhile, Kartalis *et al.*[11] found that RT yielded ADC values with better repeatability than FB and breath-hold (BH) and showed that mean ADC values in RT were higher than in FB. This differs from the findings in the present study, in which it was found that the pooled wCV of ADC was higher in RT than in FB. This discrepancy may be due to differences in equipment and sample populations. For example, Kartalis *et al.*[11] used a 1.5 T scanner and included pancreatic cancer patients, who may exhibit more motion artifacts than healthy volunteers.

However, the above finding contrasts with the study by Ma *et al.*,[21] who reported that ADC obtained from RT and BH had better accuracy and repeatability than that obtained from FB. Such discrepancies may arise from differences in study populations (e.g., healthy volunteers vs patients), imaging protocols, or analytical methodologies. These differences highlight the need for further studies to evaluate the effects of respiratory control techniques across a broader population. Therefore, although no statistically significant differences were found between the RT and FB techniques in healthy volunteers, the observed trend of lower wCV in RT, especially in D_{slow} and f (which are considered useful for pathological assessment), suggests that RT may provide more stable and useful results in future clinical applications.

Analysis of the data distribution using boxplots revealed outliers in D_{slow} and f. Specifically, one outlier was identified for D_{slow} in both the RT and FB groups, while two outliers were found for f in the RT group. These outliers may have arisen from unusual breathing motion in one subject who reported that he felt uncomforted during scan. This inconsistent breathing throughout the scanning time could lead to voxel values deviating from the mean. Additionally, signal noise, particularly in low-signal voxels or those near organ boundaries, could contribute to abnormal values. The f parameter, which reflects microvascular perfusion, is inherently more variable than parameters like D_{slow} and ADC and is thus more prone to outliers. This result aligns with the outlier found in the within-technique Bland-Altman plot.

Although RT showed a trend toward lower wCV, the Bland-Altman analysis demonstrated generally good agreement between RT and FB, with only a small number of cases showing substantial discrepancies at the global-level analysis (average from whole pancreas). Given the limited sample size, these findings do not provide sufficient evidence to conclude that either technique is superior. Further studies with larger

cohorts are needed to confirm their comparative performance.

Voxel-wise MAE analysis showed no significant differences in D_{slow}, D_{fast}, or f, except for ADC. These findings suggest minimal effects of breathing control on voxel-wise scan–rescan repeatability. However, the ADC difference may partly reflect the different fitting models, with ADC estimated using a monoexponential model and IVIM parameters estimated using Bayesian fitting. In addition, the absence of spatial registration means that MAE reflects both fitting variability and residual inter-scan motion or misregistration. Future studies should incorporate deformable image registration to distinguish fitting variability from motion-related effects.

Limitations

This study has several limitations. Firstly, the healthy volunteers did not undergo a 6-hour fasting period prior to scanning, which may have increased organ motion and affected IVIM-DWI parameter estimation. Secondly, the phantom experiments were performed at room temperature (16±3 °C) without active temperature control, deviating from the QIBA protocol (0±0.2 °C). Because ADC is temperature-dependent, phantom measurements may differ from standardized reference values; therefore, the phantom experiments were intended primarily for relative comparison of b-value schemes. Thirdly, the study included only young, healthy volunteers, limiting demographic and clinical generalizability. Fourthly, all scans were performed on a single 3.0-T Philips Ingenia system, limiting generalizability across vendors and hardware configurations. Finally, no image registration was applied between repeated scans, which may have affected voxel-level comparisons through residual respiratory displacement.

Future study

Future studies should use temperature-controlled phantom acquisitions and perfusion-sensitive phantoms or independent reference standards to validate D_{fast} and f. Image registration should be applied before voxel-wise analysis to distinguish true parameter variability from motion-related effects. Larger and more diverse patient cohorts should also be studied to establish the clinical utility and generalizability of the optimized IVIM-DWI protocol.

Conclusion

This preliminary study demonstrates the feasibility of an optimized IVIM-DWI protocol for pancreatic perfusion assessment. The 13 b-value protocol provided higher SNR and lower variability. Although wCV did not differ significantly between RT and FB, RT showed greater stability in key parameters, particularly D_{slow} and f, supporting its potential clinical utility. Further studies with larger cohorts should validate the optimized IVIM-DWI protocol against DCE-MRI to assess its accuracy for pancreatic perfusion evaluation and establish a standardized non-contrast perfusion protocol.

Ethical approval

Ethical approval was granted by the Human Research Ethics Committee of the research site (Approval Code: EC 122/2567), valid through 27 February 2026. Written informed consent was obtained from all individual participants prior to the MRI scans.

Declaration of generative AI and AI-assisted technologies

During the preparation of this manuscript, the authors used Claude (Anthropic; model Claude Sonnet 5) to assist with improving the English language only. All AI-assisted output was reviewed, verified, and edited by the authors under full human oversight, and the authors take complete responsibility for the accuracy, originality, and scientific integrity of the final manuscript.

Funding

This research was funded by the Faculty of Health Science Technology, Chulabhorn Royal Academy (Project Code: STU10767). The funding period ended on 3 April 2026.

Conflict of interest

The authors declare no conflict of interest.

CRediT authorship contribution statement

Phimphitcha Ratiphunpong: conceptualization, methodology, formal analysis, investigation, validation, writing-original draft, review and editing, project administration; **Kanyanat Sukiam**: methodology, formal analysis, investigation, data curation, writing-original draft; **Chompoonut Arsakij**: methodology, formal analysis, investigation, data curation, writing-original draft; **Nicharat Klaewsarikorn**: methodology, formal analysis, investigation, data curation, writing-original draft; **Monchai Phonlakrai**: conceptualization, methodology, formal analysis, investigation, validation, writing-original draft, review and editing, supervision, project administration, funding acquisition.

Acknowledgements

We would like to thank the Diagnostic and Interventional Radiology Departments, Chulabhorn Hospital, for research facility support.

References

[1] Smits MM, van Geenen EJM. The clinical significance of pancreatic steatosis. Nat Rev Gastroenterol Hepatol. 2011; 8(3): 169-77. doi: 10.1038/nrgastro.2011.4.

[2] Ma W, Zhang G, Ren J, et al. Quantitative parameters of intravoxel incoherent motion diffusion weighted imaging (IVIM-DWI): potential application in predicting

pathological grades of pancreatic ductal adenocarcinoma. Quant Imaging Med Surg. 2018; 8(3): 301-10. doi: 10.21037/qims.2018.04.08.

[3] Pham VT, Ciccaglione M, Ramirez DG, Benninger RKP. Ultrasound imaging of pancreatic perfusion dynamics predicts therapeutic prevention of diabetes in preclinical models of type 1 diabetes. Ultrasound Med Biol. 2022; 48(7):1336-47. doi: 10.1016/j.ultrasmedbio.2022.03.013.

[4] Kubo S, Yamamoto K, Magata Y, et al. Assessment of pancreatic blood flow with positron emission tomography and oxygen-15 water. Ann Nucl Med. 1991; 5(4): 133-8. doi: 10.1007/BF03164627.

[5] Zaborienė I, Barauskas G, Gulbinas A, Ignatavičius P, Lukoševičius S, Žvinienė K. Dynamic perfusion computed tomography: a promising tool to diagnose pancreatic ductal adenocarcinoma. Open Med (Wars). 2021; 16(1): 284-92.doi: 10.1515/med-2021-0228.

[6] Hu R, Yang H, Chen Y, et al. Dynamic contrast-enhanced magnetic resonance imaging for measuring pancreatic perfusion in acute pancreatitis: a preliminary study. Acad Radiol. 2019; 26(12): 1641-9.doi: 10.1016/j.acra.2019.02.007.

[7] Le Bihan D, Breton E, Lallemand D, Aubin ML, Vignaud J, Laval-Jeantet M. Separation of diffusion and perfusion in intravoxel incoherent motion MR imaging. Radiology. 1988; 168(2): 497-505. doi: 10.1148/radiology.168.2.3393671.

[8] Dixon WT. Separation of diffusion and perfusion in intravoxel incoherent motion magnetic resonance imaging: a modest proposal with tremendous potential. Radiology. 1988; 168(2): 566-7. doi: 10.1148/radiology.168.2.3393682.

[9] Szubert-Franczak AE, Naduk-Ostrowska M, Pasicz K, Podgórska J, Skrzyński W, Cieszanowski A. Intravoxel incoherent motion magnetic resonance imaging: basic principles and clinical applications. Pol J Radiol. 2020; 85:e624-35. doi: 10.5114/pjr.2020.101476.

[10] Nadolska K, Białecka A, Zawada E, Kazimierczak W, Serafin Z. Analysis of intravoxel incoherent motion perfusion fraction improves detection of pancreatic ductal adenocarcinoma. Diagnostics (Basel). 2024; 14(6): 571.doi: 10.3390/diagnostics14060571.

[11] Kartalis N, Loizou L, Edsborg N, Segersvärd R, Albiin N. Optimising diffusion-weighted magnetic resonance imaging for demonstrating pancreatic cancer: a comparison of respiratory-triggered, free-breathing and breath-hold techniques.Eur Radiol. 2012; 22(10): 2186-92. doi: 10.1007/s00330-012-2469-3.

[12] Bäuchle TA, Stuprich CM, Loh M, Nagel AM, Uder M, Laun FB. Influence of magnetic field strength on intravoxel incoherent motion parameters in diffusion MRI of the calf. Tomography. 2024; 10(5): 773-88.doi: 10.3390/tomography10050059.

[13] Grimbergen G, Eijkelenkamp H, Heerkens HD, Raaymakers BW, Intven MPW, Meijer GJ. Intrafraction pancreatic tumor motion patterns during ungated magnetic resonance guided radiotherapy with an abdominal corset. Phys Imaging Radiat Oncol. 2022; 21: 1-5. doi: 10.1016/j.phro.2021.12.001.

[14] Gurney-Champion OJ, Froeling M, Klaassen R, et al. Minimizing the acquisition time for intravoxel incoherent motion magnetic resonance imaging acquisitions in the liver and pancreas. Invest Radiol. 2016; 51(4): 211-20.doi: 10.1097/RLI.0000000000000225.

[15] Gustafsson O, Brynolfsson P, Galldin T, Cedersund G, Wirestam R. Impact of prior distributions and central tendency measures on Bayesian intravoxel incoherent motion model fitting. Magn Reson Med. 2018; 79(3): 1674-83.doi: 10.1002/mrm.26783.

[16] Jalnefjord O, Thörnqvist S, Wirestam R. Comparison of methods for estimation of the intravoxel incoherent motion (IVIM) diffusion coefficient (D) and perfusion fraction (f). MAGMA. 2018; 31(5): 715-23. doi: 10.1007/s10334-018-0697-5.

[17] Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement.Lancet. 1986; 1(8476): 307-10. doi: 10.1016/S0140-6736(86)90837-8.

[18] Ogura A, Hatano I, Osakabe K, Yamaguchi N, Koyama D, Watanabe H. Importance of fractional b-value for calculating apparent diffusion coefficient in DWI. AJR Am J Roentgenol. 2016; 207(6): 1239-43. doi: 10.2214/AJR.15.15945.

[19] Freiman M, Voss SD, Mulkern RV, Perez-Rossello JM, Callahan MJ, Warfield SK. In vivo assessment of optimal b-value range for perfusion-insensitive apparent diffusion coefficient imaging. Med Phys. 2012; 39(8): 4829-32.doi: 10.1118/1.4736516.

[20] Liu X, Wang YF, Qi XH, et al. Reproducibility study of intravoxel incoherent motion and apparent diffusion coefficient parameters in normal pancreas. World J Gastrointest Surg. 2024; 16(7): 2031-9. doi: 10.4240/wjgs.v16.i7.2031.

[21] Ma C, Wang J, Li YJ, et al. Comparisons of image quality and ADCs in breath-hold, respiratory-triggered and free-breathing DWI of pancreas at 3-T. Open J Radiol. 2014; 4(4): 279-92. doi: 10.4236/ojrad.2014.44037.