
The role of diffusion weighted imaging in differentiation of solid renal masses



Master Thesis Technical Medicine

by

Ilse Manon Spenkelink

**UNIVERSITY
OF TWENTE.**

Radboudumc

August 29, 2019

Graduation committee

Chairmen

Prof. dr. J.J. Fütterer
*Department of Radiology, Radboud University medical center,
Nijmegen, The Netherlands*

Medical supervisors

Prof. dr. J.J. Fütterer
*Department of Radiology, Radboud University medical center,
Nijmegen, The Netherlands*

Drs. T.J. van Oostenbrugge
*Department of Urology, Radboud University medical center,
Nijmegen, The Netherlands*

Technological supervisors

Prof. dr. ir. C.H. Slump
*MIRA Institute for Biomedical Engineering and Technical
medicine, University of Twente, Enschede, The Netherlands*

Dr. C.G. Overduin
*Department of Radiology, Radboud University medical center,
Nijmegen, The Netherlands*

Process supervisor

Dr. M. Groenier
*Department of Technical Medicine, University of Twente,
Enschede, The Netherlands*

External member

A. Christenhusz, MSc
*Department of Magnetic Detection & Imaging, University of
Twente, Enschede, The Netherlands*

Preface and acknowledgments

Before you lies the final product of my thesis, written to obtain the Master's degree of Technical Medicine. In this thesis, I present the results of my research concerning differentiation of solid renal masses, a research conducted between October 2018 and August 2019 at the Radiology Department of the Radboudumc Nijmegen.

The past 10 months have been full of challenges and opportunities. I have learned a great deal and therefore I would like to thank my supervisors. I appreciate your support, advice and inspiration, in terms of conducting research, as well as in personal development. You answered my questions and helped me in putting things into perspective. You all have contributed to my Master's, each in your own way. Thank you for that.

I hope you enjoy reading this thesis.

Ilse Spenkelink

Nijmegen, August 29, 2019

Abstract

Objective: The purpose of this study was to determine the utility of diffusion weighted imaging (DWI) parameters in differentiating solid renal masses.

Materials and methods: The performance of mono-exponential DWI-parameters combined with patient characteristics in differentiating oncocytomas (OCs) and renal cell carcinomas (RCCs) was evaluated. DWI-imaging was performed in 39 patients (7 OCs and 32 RCCs) using three b-values. Whole-tumor apparent diffusion coefficient (*ADC*) maps and histogram parameters were calculated. The parameters were compared between both groups by the Mann-Whitney test. Significantly different parameters were used in a predication model. The diagnostic value of this model was evaluated using receiver operating characteristic (ROC) curve analyses.

Hereafter, simulations were performed to obtain optimized b-value distributions and to compare the trust-region reflective (TRR) and Bayesian probability (BP) algorithms for fitting of the intravoxel incoherent motion (IVIM) and stretched exponential DWI-models. Optimized b-value distributions were acquired by minimizing the parameter estimation error. The results were verified in four healthy volunteers. This was followed by a prospective study to determine the role of both multi-exponential models in differentiating renal tumors. Up until now, DWI-images were obtained in five patients (3 clear cell RCC (ccRCC) and 2 papillary RCC (pRCC)). Again, whole-tumor DWI and histogram parameter analyses were performed. Additionally tumor induced changes in histogram parameters were analyzed.

Results: Standard deviation and entropy of *ADC*, tumor volume, and sex, were statistically significantly different between OCs and RCCs ($p < 0.05$). The area under the ROC curve was 0.93, with a maximum sensitivity and specificity of 86% and 84%, respectively.

The simulations provided optimized b-value distributions containing eight b-values. In simulations and in vivo, the BP algorithms provided more robust parameter estimations compared to the TRR algorithm, for the IVIM model. In the stretched exponential model, there was no pronounced difference.

The 75th and 95th percentile of the true tissue diffusion parameter (D_t) and the mode of the distributed diffusion coefficient (*DDC*) showed no overlap between ccRCC and pRCC. Furthermore, no overlap was found for several tumor induced histogram parameters changes: pseudodiffusion fraction (*f*), D_t , *DDC* and the diffusion heterogeneity index (α).

Conclusion: Histogram analysis of DWI-parameters aid differentiation of benign and malignant renal masses. Addition of patient characteristics might improve this differentiation. Multi-exponential DWI-parameters can be robustly estimated using eight b-values. These parameters may provide a better description of tissue diffusion characteristics and could therefore improve differentiation of renal tumors. Lastly, tumor induced differences in histogram parameters may be valuable for tumor differentiation.

Table of contents

Graduation committee	ii
Preface and acknowledgments.....	iii
Abstract	iv
List of abbreviations	vii
List of Figures	viii
List of Tables	ix
CHAPTER 1 – General Introduction	1
1.1 Background	1
1.1.1 Kidney tumors	1
1.1.2 Differentiation of solid renal masses	1
1.1.3 Diffusion-weighted imaging.....	1
1.2 Objectives.....	3
1.3 Thesis outline	3
CHAPTER 2 – Utility of a prediction model comprising diffusion-weighted imaging parameters and patient characteristics in differentiation of benign and malignant renal tumors.....	4
2.1 Introduction	4
2.2 Materials and methods	4
2.2.1 Patients.....	4
2.2.2 Imaging protocol	5
2.2.3 Image analysis	5
2.2.4 Pathological assessment.....	5
2.2.5 Statistical analysis.....	5
2.3 Results	7
2.3.1 Patient characteristics.....	7
2.3.2 Renal tumor comparison	7
2.3.3 Incremental use of patient demographics.....	8
2.4 Discussion.....	9
2.5 Conclusion.....	10
CHAPTER 3 – Evaluation of <i>b</i> -value distributions and fitting algorithms in estimating multi-exponential diffusion-weighted imaging parameters	11
3.1 Introduction.....	11
3.1.1 Diffusion-weighted imaging models	11
3.1.2 <i>b</i> -value selection	11
3.1.3 Data fitting strategies.....	12
3.1.4 Diffusion weighted magnetic resonance imaging sequence	12
3.1.5 Research aim	13
3.2 Methods.....	13
3.2.1 Simulations	13

3.2.2 Algorithm optimization criteria	14
3.2.3 Combined b -value distribution selection.....	14
3.2.4 Human subject validation	14
3.3 Results.....	15
3.3.1 Simulations	15
3.3.2 In vivo measurements.....	15
3.4 Discussion	21
3.4.1 Simulations	21
3.4.2 In vivo validation	21
3.5 Conclusion.....	22
CHAPTER 4 – Diagnostic utility of multi-exponential diffusion-weighted imaging parameters in differentiating solid renal tumors.....	24
4.1 Introduction	24
4.2 Methods	24
4.2.1 Patients	24
4.2.2 MR Imaging.....	24
4.2.2 Image analysis	25
4.2.3 Histologic analysis	25
4.2.4 Statistical analysis.....	25
4.3 Results	25
4.3.1 Patient characteristics.....	25
4.3.2 Subtype comparison	25
4.3.3 Statistical analysis.....	27
4.4 Discussion.....	27
4.5 Conclusion.....	30
CHAPTER 5 – General conclusion, discussion and future perspectives	31
References	32
Appendix A Design parameters for computer simulations	37
A.1 b -value distribution	37
A.2 Data noise	37
A.3 Fitting algorithms.....	37
A.4 Data fitting	38
Appendix B Parameter estimation errors.....	39
Appendix C Parameter errors and standard deviations for combinations of multi-exponential DWI-models	46
Appendix D Comparison of central tendency measures of the Bayesian probability algorithm in multi-exponential DWI-parameter histogram parameters.....	47
Appendix E Multi-exponential DWI-parameter histogram analysis of tumor induced differences.....	52

List of abbreviations

α	Diffusion heterogeneity index
ADC	Apparent diffusion coefficient
AML	Angiomyolipoma
fpAML	Fat-poor angiomyolipoma
BP	Bayesian probability
<i>DDC</i>	Distributed diffusion coefficient
D_p	Pseudodiffusion
D_t	True tissue diffusion
DWI	Diffusion-weighted imaging
f	Pseudodiffusion fraction
HASTE	Half-Fourier-acquired single-shot turbo spin echo
IVIM	Intravoxel incoherent motion
OC	Oncocytoma
RCC	Renal cell carcinoma
ccRCC	Clear cell renal cell carcinoma
chRCC	Chromophobe renal cell carcinoma
pRCC	Papillary renal cell carcinoma
SMS	Simultaneous multislice
SS-EPI	Single-shot echo planar imaging
TRR	Trust-region reflective

List of Figures

Figure 1. The principle of diffusion-weighted imaging	2
Figure 2. Overview of histogram parameters	3
Figure 3. Image analysis workflow	5
Figure 4. MRI image and diffusion-weighted imaging analysis of an oncocytoma	8
Figure 5. ROC-curves of classification model to predict oncocytoma	10
Figure 6. Histograms of b -value distributions obtained by computer simulations	16
Figure 7. Diffusion-weighted imaging parameter estimation errors and standard deviations	17
Figure 8. Multi-exponential diffusion-weighted imaging parameter maps	19
Figure 9. Parameter maps of multi-exponential diffusion-weighted imaging parameters in a papillary RCC	26
Figure A1. Segmented fitting method	37
Figure B1 Fitting errors of apparent diffusion coefficient	39
Figure B2 Fitting errors of the pseudodiffusion fraction	40
Figure B3 Fitting errors in the pseudodiffusion parameter	41
Figure B4 Fitting errors in the true tissue diffusion parameter	42
Figure B5 Fitting errors in the distributed diffusion coefficient	43
Figure B6 Fitting errors in the diffusion heterogeneity index	44
Figure B7 Overview of fitting errors in parameters that reflect tissue diffusion	45
Figure E1 Histogram parameters of tumor induced differences	47

List of Tables

Table 1. Patient demographics and tumor characteristics.....	6
Table 2. MRI scan parameters	7
Table 3. Demographical and histogram apparent diffusion coefficient parameters.....	7
Table 4. MRI scan parameters in healthy volunteers	13
Table 5. Optimized b -value distributions for separate multi-exponential DWI-models.....	16
Table 6. Optimized b -value distributions for combined multi-exponential DWI-models	18
Table 7. Multi-exponential DWI-parameters in healthy volunteers	18
Table 8. MRI scan parameters for multi-exponential DWI-parameter estimation	26
Table 9. Patient demographics and tumor characteristics.....	27
Table 10. Multi-exponential histogram parameters in renal cell carcinoma	28
Table 11. Tumor induced changes in histogram parameters in renal cell carcinoma	29
Table A1 Overview of diffusion-weighted imaging parameters used for simulations.....	38
Table C1 DWI parameter estimation errors of the optimized b -value distribution.....	46
Table C2 Parameter errors and standard deviation for all combinations DWI-models and fitting algorithms.....	46

CHAPTER 1 – General Introduction

1.1 Background

1.1.1 Kidney tumors

Over the past two decades, the incidence of renal masses has been increasing. This can be explained by the increase of incidental findings on diagnostic ultrasound, computed tomography and magnetic resonance imaging examinations performed for other reasons than to detect renal tumors. Small renal masses (≤ 4 cm) account for almost half of the newly found renal tumors. Present-day imaging techniques are successful in identifying these lesions. However, reliable preoperative differentiation between malignant and benign solid renal masses remains challenging.

Several genetic factors, the male sex and an age between the sixth and eighth decade are associated with an increased risk of developing malignant renal tumors. Furthermore, lifestyle factors that affect body mass index (BMI), hypertension and smoking are identified as risk factors.[1]

Accounting for 90% of the renal malignancies, the most common malignant tumor type is renal cell carcinoma (RCC). RCCs are subdivided into fourteen subtypes according to the 2016 World Health Organization (WHO) renal tumor classification.[2] The three most frequent subtypes, comprising more than 90% of the RCCs, are clear cell RCC (ccRCC, 75%), papillary RCC (pRCC, 10-15%) and chromophobe RCC (chRCC, 5%). The subtypes differ in pathological and clinical features, including growth rate and propensity to metastasize. Furthermore, the subtypes show differences in prognosis, with 5-year cancer specific survival rates of 68.9%, 87.4%, and 86.7% for ccRCC, pRCC, and chRCC, respectively.[3]

The most common benign renal tumors are angiomyolipomas (AML) and oncocytomas (OC), comprising 44% and 35% of the resected benign solid renal masses, respectively.[4] In general, these tumors have a slow and consistent growth rate and do not metastasize. Therefore, active surveillance is an appropriate option for most of these benign tumors.[1]

1.1.2 Differentiation of solid renal masses

Due to differences in prognosis and treatment options recommended by the national comprehensive cancer network (NCCN), reliable preoperative differentiation between benign and malignant renal masses, as well as between the RCC subtypes, is important. This differentiation is impeded since benign renal tumors can have similar imaging features to RCC. OCs and chRCC originate from the same progenitor cell. As a consequence, histological tumor properties overlap, which complicates differentiation. Additionally, differentiation between OCs and ccRCC is considered difficult, because imaging characteristics and enhancement patterns are not fully tumor type specific.[5] Furthermore central scars that are seen in OCs can mimic necrosis within RCCs and a small fraction of RCCs contain scars.[6] AMLs can often be distinguished from other renal tumors by their fat component. However, 5% of the AMLs contain no or only a small amount of fat, which makes it difficult to differentiate them from RCCs using CT and MRI. These AMLs are referred to as fat-poor AML (fpAML).[7] fpAMLs show hyperattenuation and hypointensity on unenhanced CT and T2-weighted MRI scans, respectively. These findings are also seen in RCCs, in particular in pRCCs.[8]

Nowadays, renal tumors are primarily treated by (partial) nephrectomy. Recent studies show that 20% of the resected small renal masses are benign at histopathological analysis. Preoperative percutaneous biopsies can be used for differentiation of the tumor types. High sensitivities (86-100%) and specificities (98-100%) are reported for renal mass biopsy. However, in 22% of the cases, the biopsies are non-diagnostic as a result of sampling error.[9] CT and MRI provide 3D information about the lesions and therefore overcome sampling error. When sensitivity and specificity for tumor differentiation can be increased using these imaging techniques, a more accurate and non-invasive lesion characterization can be provided. Thereby, the rate of unnecessary surgery and, consequently, patient morbidity and costs will be reduced.

1.1.3 Diffusion-weighted imaging

Prior studies have shown that diffusion-weighted imaging (DWI) is a promising MRI technique for the evaluation of renal lesions.[10]–[13] DWI is performed by using additional pairs of gradients during the MRI sequence. The degree of diffusion weighting depends on the magnitude and duration of the gradients and the time between them. (Figure 1) This diffusion weighting is referred to as the *b*-value. Images acquired with a high *b*-value are very

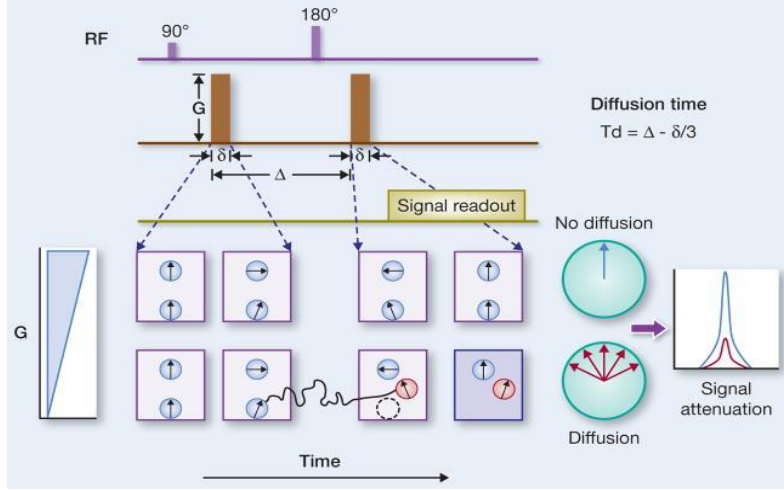


Figure 1. Diffusion weighted images are acquired by using additional pairs of gradient pulses with magnitude G and duration δ . During the first pulse, protons are phase-encoded, depending on their spatial location. The second gradient, which realigns stationary spin phases, is applied after delay-time Δ . Imperfect refocusing occurs when water molecules diffused to another spatial location along the gradient field direction. This results in phase dispersion, which causes attenuation of the magnitude signal. T_d : effective diffusion time. Adapted from “Role of diffusion weighted MR imaging in characterization of focal kidney and upper urinary tract lesions”, by ‘D. M. Sobh et al.

sensitive to motion. As a result, most tissues will have lost their signal due to molecular motion, while restricted lesions will be hyperintense. When acquiring DW-images with diverse b -values, DWI-parameters can be calculated.

Strictly speaking, diffusion can be free, hindered or restricted. Free diffusion portrays the thermally driven, microscopic random motion of water molecules in absence of obstacles, also referred to as Brownian motion. In this case, the displacement distribution follows a Gaussian distribution. In renal tissue, water molecules within the extracellular space interact with cell membranes and macromolecules. Therefore, diffusion of water molecules is hindered. When water molecules are trapped within an enclosed compartment, the diffusion distance is confined by the dimensions of the compartment. This is called restricted diffusion. However, in DWI literature the term ‘restricted diffusion’ is used for a decrease in diffusion distance caused by restricted diffusion as well as hindered diffusion. This definition will be used in the current thesis as well.

Each tissue has its characteristic cellular architecture and distribution of intra- and extracellular space and therefore its specific diffusion characteristics. Tumor growth induces changes in tissue structure which thus results in changes in diffusion parameters. In solid renal tumors, the high cellular density is believed to cause diffusion restriction.[12], [13] Several studies have investigated the role of DWI in renal mass characterization using a mono-exponential model.[11], [14]–[18] This model describes the relationship between the measured signal intensity using a certain b -value (S_b) and the apparent diffusion coefficient (ADC), which reflects tissue diffusivity:

$$S_b = S_0 e^{-b \cdot ADC} \quad (1)$$

However, ADC values are influenced by the selected b -values.[18] The use of lower maximum b -values results in higher ADC ($ADC = 2.74 \pm 0.58 \text{ mm}^2/\text{s}$ using $b = 0$ and $b = 100 \text{ s/mm}^2$ compared to $1.66 \pm 0.51 \text{ mm}^2/\text{s}$ using $b = 0$ and $b = 1000 \text{ s/mm}^2$ in malignant renal lesions). Furthermore, the broad range in mean ADC values causes a significant overlap between RCC subtypes and between malignant and benign solid renal masses. Therefore, the ADC value on its own is no reliable predictor for renal malignancy.

1.1.3.1 Diffusion parameter histogram analysis

Quantification of differences in the degree of heterogeneity and histologic structure between renal tumor subtypes may improve discrimination of the solid renal masses.[19], [20] Information about the distribution of diffusion parameters can be obtained by whole-tumor histogram analysis. First-order texture metrics that can be derived from the histograms of diffusion parameters are, amongst others, mean, skewness, kurtosis, entropy and percentiles. Skewness resembles the degree of asymmetry of the distribution of values, kurtosis is the degree of

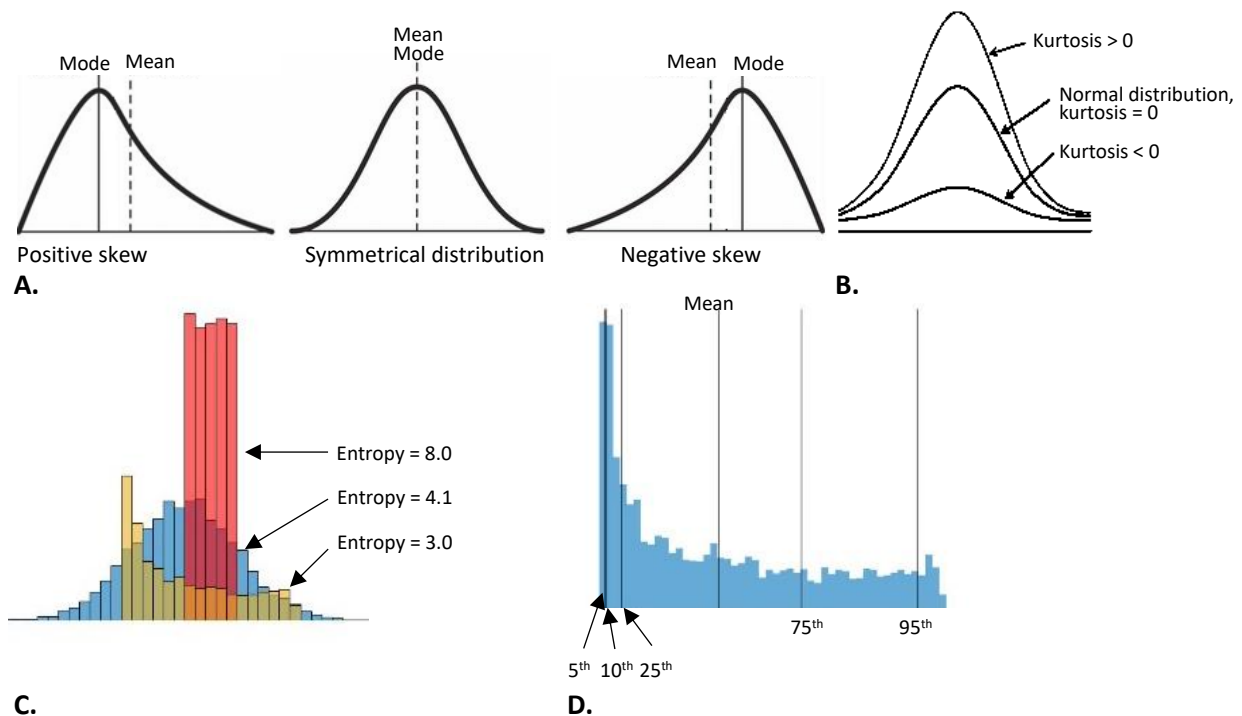


Figure 2. Overview of histogram parameters. Examples of a symmetrical, positively and negatively skewed distribution are shown in A. Examples of a normal distribution and distributions with an positive and negative kurtosis are shown in B. Distributions with different measures of entropy are shown in C. In D, a distribution is shown with its 5th, 10th, 25th, 75th and 95th percentiles

peakedness, and entropy describes the variation and unpredictability of individual values in the histogram, which is maximal at an uniform distribution. An overview of the histogram parameters is shown in Figure 2.

1.2 Objectives

To improve treatment decisions, it is essential to distinguish malignant from benign renal tumors, preoperatively. However, because of the differences in prognosis, it is also important to discriminate the RCC subtypes. Thus far, it is not possible to predict the renal tumor type in a non-invasive and reliable way. Therefore, the main research question of this thesis is: What is the role of diffusion-weighted imaging in the differentiation of solid renal masses?

To address this question, the following sub-questions were formulated:

1. To what extent can the combination of mono-exponential diffusion-weighted imaging parameters and patient characteristics differentiate oncocytoma from renal cell carcinoma?
2. What *b*-value distribution and fitting algorithms provide a robust calculation of multi-exponential diffusion-weighted imaging parameters?
3. What is the ability of multi-exponential diffusion-weighted imaging parameters in subtyping of renal tumors?

1.3 Thesis outline

A general introduction on the role of DWI in differentiating solid renal masses was provided and the main objective was formulated. (*Chapter 1*) This is followed by a prospective observational study in which the performance of a prediction model for renal tumor benignity based on *ADC* values, diffusion histogram parameters and patient characteristics was assessed. The main objective was to differentiate OC from RCC. (*Chapter 2*) Hereafter, the MRI scan protocol was optimized for multi-exponential DWI algorithms. The goal was to minimize fitting errors caused by *b*-value sampling and the applied fitting algorithm, and to obtain an adequate scan protocol, taking the scan duration into account. (*Chapter 3*) Subsequently, this protocol was used in a prospective observational study in which the role of intravoxel incoherent motion (IVIM) and stretched exponential parameters in differentiating renal tumor types was explored. The goal was to evaluate which clinical, DWI and histogram parameters are beneficial in a prediction model. (*Chapter 4*) This thesis concludes with a general discussion and conclusion. An outline for future research is provided. (*Chapter 5*)

CHAPTER 2 – Utility of a prediction model comprising diffusion-weighted imaging parameters and patient characteristics in differentiation of benign and malignant renal tumors

2.1 Introduction

RCC comprises 85% of all renal cancers. Accounting for 3% and 5% of all diagnosed tumors, RCC is the sixth most common malignancy in males and the tenth in females.[21] The majority of renal lesions suspected for RCC are incidental findings on imaging performed for other reasons than to detect renal tumors.[22] Contrast-enhanced CT and MRI are used to differentiate lesions that are suspicious for malignancy within this group of renal tumors.[1] With sensitivities of 79.7% and 88.1%, and specificities of 44.4% and 33.3%, respectively, these cross-sectional imaging techniques cannot reliably distinguish RCC from several benign renal tumors, specifically OC.[8], [23], [24] The inadequacy in discrimination is due to the appearance of OC as solitary, well-demarcated, hypervascular, contrast-enhancing renal cortical tumors on imaging, which mimic the appearance of RCC.[25] In recent studies regarding partial nephrectomy, up to 20% of all resected lesions suspected for RCC are shown to be benign, with the majority of these lesions being OC.[26]

The 5-year survival rates of patients with RCC range from 9 to 81%, with a mean of 55.9%, depending on the TNM stage, nuclear grade and RCC subtype. Contrarily, OCs have no malignant potential and only sporadically show renal vein invasion, local recurrence after surgery or distant metastases. Even in case of dissemination, the prognosis is excellent.[27] Therefore, active surveillance is considered a proper management alternative for histologically proven OCs, especially in case of smaller tumors.[1], [28] Thus, invasive diagnostics, such as percutaneous renal biopsies, and superfluous surgery can be avoided if OCs can be noninvasively discriminated from RCC on imaging.

By assessing the restriction of free water movement, diffusion-weighted (DW) MRI can provide information about changes in local tissue structure and the increased cellularity within tumors.[29] This imaging technique has been increasingly studied in evaluating renal tumors, over the last decade.[30] In renal tumors, *ADC* values of manually selected regions of interest (ROIs) are lower for RCC compared to OC.[31] Recent studies suggest that whole tumor DW-MRI parameters, such as the *ADC* value distribution within the tumor (whole-tumor *ADC* histogram analysis), may improve small renal mass characterization.[19], [29], [32], [33] Apart from diffusion parameters, strong predictors for renal tumor malignancy are tumor size and male sex.[34] We hypothesize that a composite classification model that combines DWI-parameters and demographic parameters can improve discrimination between RCC and OC.

The purpose of this research was to assess the ability to discriminate OC from RCC based on whole tumor *ADC* parameters and patient demographic characteristics.

2.2 Materials and methods

2.2.1 Patients

This prospective study was compliant with Health Insurance Portability and Accountability Act, approved by our institutional review board. Written informed consent was obtained from each eligible patient before enrollment. Between October 2014 and June 2018, a total of 48 consecutive patients were included in the study at the Radboud university medical center. Inclusion criteria were: patients diagnosed with a renal tumor based on previous ultrasonography, CT or MRI examinations; at least 18 years of age; scheduled for a renal mass biopsy or partial or radical nephrectomy. Exclusion criteria were: contraindications for MR imaging (non-MR compatible metal device/foreign bodies, claustrophobia); an active renal or peri-renal infection; prior treatment for renal malignancies; artifacts in DWI-images impeding reliable parameter calculation; and no histopathological confirmation of renal tumor type. In case of presence of multiple tumors in one patient, only the largest tumor was used for analysis due to the expected similarity in histology of the other tumors.

Patients' renal medical history, as well as the information about diagnoses, biopsies, and surgical procedures were collected.(Table 1) Pre-operative chronic kidney disease (CKD) stage based on estimated glomerular filtration rate (eGFR) was assessed according to standard clinical practice.[35] Demographic data concerning patients' age, sex and body mass index (BMI) were also collected.[34]

2.2.2 Imaging protocol

MR imaging was performed within 21 days prior to surgery or biopsy using 3T MRI system (Magnetom Trio, Siemens, Erlangen, Germany) and a combination of a 32-channel receiver coil and phased array body surface coil. Patients were positioned in a feet-first supine position. The scanning protocol included breath-hold anatomical T2-weighted multi-slice half-Fourier-acquired single-shot turbo spin echo (HASTE) sequence in axial and coronal directions. DWI was performed using a respiratory triggered coronal single-shot-echo-planar imaging (SS-EPI) sequence with b -values of 50, 800 and 1400 s/mm² and diffusion gradients applied in three directions. (Table 2)

2.2.3 Image analysis

DICOM images were exported to an offline work-station for processing. Respiratory motion artifacts in DW images with distinct b -values were minimized by semi-automatic co-registration (FireVoxel, CAI²R, New York University, NY) performed by one observer (IS). Registration was done separately for each kidney (with and without tumor), resulting in two datasets for each patient. Next, two observers blinded for clinical and histopathological information (IS and MA, with 1 and 3 years of experience in abdominal imaging, respectively), independently drew regions of interest (ROIs) on both registered DWI datasets. Both readers were supervised by a radiologist (JF, 15 years of experience in abdominal radiology). In tumor-bearing kidneys, the ROI covered the entire tumor on all slices where tumor was visible, including necrotic and cystic parts. The tumor volumes were measured. The ROI drawn in the contralateral kidney dataset covered a part of 1-3 cm³ of renal cortex.

To reduce partial volume effects, edge voxels of the tumor ROI were removed using an erosion method. The mono-exponential DWI model (1) was fitted voxel-wise to the measured signal intensities, using a nonlinear least-squares method. (MATLAB R2019a, MathWorks; Natick; MA, USA). The ADC values of the tumor and renal cortex were then used for histogram analysis.

Outliers of ADC values in the histogram were detected and removed according to the interquartile range (IQR) method, using an outlier rejection factor of 2.2, to reduce their effect on ADC histogram parameters.[36] The following ADC histogram parameters were derived: mean, mode, standard deviation (SD), 5th percentile, 10th percentile, 25th percentile, 75th percentile, 95th percentile, entropy, skewness, and kurtosis. (Figure 3 and 4) Inter-observer agreement for the tumor volume was tested for ten randomly selected cases.

2.2.4 Pathological assessment

Percutaneous renal biopsy and surgical pathology specimens obtained from solid tumor regions were fixed in formalin, embedded in paraffin and sliced into four-micron thick sections. H&E-stainings were used for pathological analysis. Histopathological diagnoses were made by board-certified pathologists, according to WHO classification on renal tumors.[2]

2.2.5 Statistical analysis

For descriptive analysis, median (range) and mean (\pm SD) values were calculated. Interobserver agreement between the tumor volumes obtained by the two readers was tested using an intraclass correlation coefficient (ICC). ICC >0.80 was interpreted as an excellent agreement. If the ICC was >0.80 the data from one observer (IS) were used, otherwise data of both observers were analyzed independently.

The Mann-Whitney test was used to compare the ADC histogram parameters and patients' age and BMI between both patient groups. The differences between the groups by patient sex were assessed using the Chi-square test.

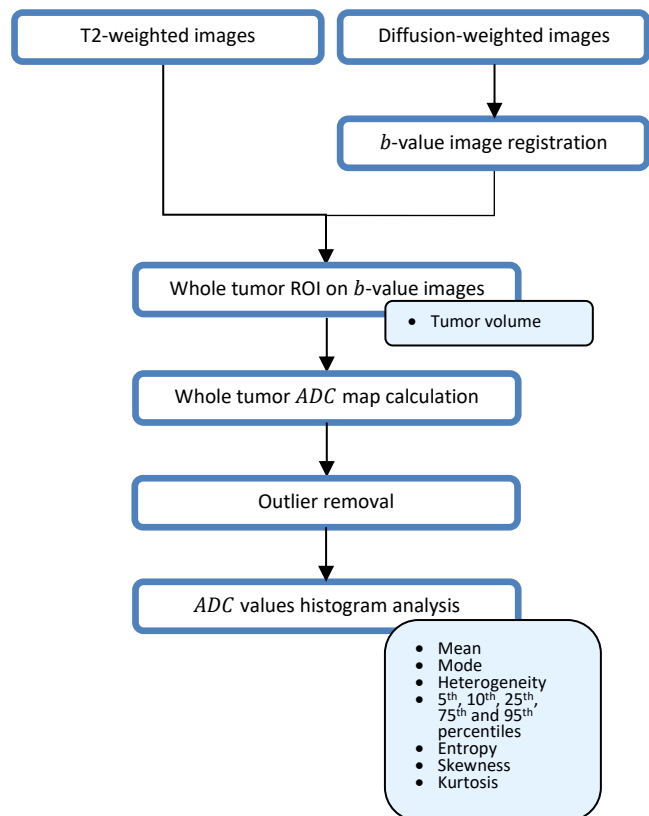


Figure 3. Image analysis workflow.

To discriminate OC from RCC, a logistic regression model was created using parameters that were significantly different between the two tumor types. Collinearity of model parameters was tested using the variance inflation factor (*VIF*), defined as $VIF = \frac{1}{1-R^2}$, where R^2 is the coefficient of determination between a given predictor (parameter) and all other predictors. A *VIF* > 4 was considered as presence of multicollinearity, and such parameters were excluded from the model.

First, logistic regression analysis for predicting OC was performed using only *ADC* histogram parameters that showed statistically significant differences between the two patient groups. A second model was constructed by adding the demographic parameters. A Receiver Operating Characteristic (ROC) curve for both regression models was created and the area under the curve (AUC) was calculated. The optimal sensitivity and specificity of the ROC were selected by maximizing the Youden's index ($J = \text{sensitivity} + \text{specificity} - 1$).

A separate analysis was performed to evaluate inpatient *ADC* histogram parameter differences between the tumor and non-tumor renal cortical tissue. A Mann-Whitney test was used to compare the tumor-induced differences between the benign and malignant tumor groups. Statistically significantly different parameters were used for logistic regression analysis to predict benign nature. The ROC-AUC was calculated for this regression model. Analyses were done using Statistical Package for the Social Sciences version 25.0 (SPSS, IBM, Armonk, New York, USA). For all tests, a p-value ≤ 0.05 was considered statistically significant.

Table 1. Patient demographics and tumor characteristics

	Renal cell carcinoma (n=32 patients)	Oncocytoma (n=7 patients)
Patient age (y)	65 (38-79)	65 (28-75)
Female	5 (16)	4 (57)
Body mass index (kg/m ²)	27.8 (22.0-36.7)	26.9 (20.7-35.2)
Tumor volume (cm ³)	262 (6-1261)	48 (3-169)
Affected side		
Left	16 (50.0)	2 (28.6)
Right	16 (50.0)	5 (71.4)
Type of procedure		
Laparoscopic partial nephrectomy	7 (21.9)	4 (57.1)
Open partial nephrectomy	1 (3.1)	0
Laparoscopic radical nephrectomy	9 (28.1)	0
Open radical nephrectomy	18 (46.9)	0
Percutaneous biopsy	0	3 (42.9)
T-stage		
T1	14 (43.8)	6 (85.7)
T2	1 (3.1)	1 (14.3)
T3	16 (50.0)	0
T4	1 (3.1)	0
Histological subtype RCC		NA
Clear cell	25 (78.1)	
Papillary	3 (9.4)	
Chromophobe	2 (6.3)	
Papillary clear cell	1 (3.1)	
Mucinous tubular and spindle cell	1 (3.1)	
Nuclear grading (Fuhrman)		NA
I	1 (3.1)	
II	10 (31.3)	
III	11 (34.4)	
IV	7 (21.9)	
Unknown	3 (9.4)	

Data are presented as mean (range) or No. (%); NA = not applicable

Table 2. Magnetic resonance imaging parameters

Imaging parameter	T2WI Axial	T2WI Coronal	DWI
TR/TE, ms	2000/92	2000/103	2800/77
Flip angle, degrees	150	140	90
No. of slices	30	30	35
Slice thickness, mm	4	5	4
Field of view, mm	450 x 450	400 x 400	380 x 309
Matrix size	320 x 256	320 x 256	192 x 154
Voxel size, mm x mm	1.4 x 1.8	1.3 x 1.6	2.0 x 2.5
<i>b</i> -values, s/mm ²	NA	NA	50/800/1400

TR = repetition time; TE = echo time, T2WI = T2-weighted magnetic resonance imaging; DWI = diffusion weighted magnetic resonance imaging; NA = not applicable.

2.3 Results

2.3.1 Patient characteristics

MRI and histological tumor analysis were performed in 48 patients. Data of 9 patients were excluded for the following reasons: severe imaging artifacts within the tumor (n=2); presurgical tumor embolization (n=1); cystic tumors (n=3); and absence of histologically confirmed renal tumor (n=3). The final cohort included 39 patients (9 females, mean age 65 years (38-79)) with 39 lesions: 32 malignant RCCs (25 clear cell, 2 chromophobe, 3 papillary, 1 clear cell papillary, and 1 mucinous tubular and spindle cell RCC), and 7 OCs. (Table 1) Interobserver agreement for tumor volume was excellent (ICC, 0.98). Therefore, analyses were done using the ROIs drawn by one reader (IS).

2.3.2 Renal tumor comparison

The mean tumor volumes for RCC are 242 cm³ (median: 132.6, range 5-1261 cm³) and for OC 48 cm³ (median: 40.9, range 3-169 cm³). The tumor volumes ($p=0.012$), and the SD ($p=0.008$) and entropy ($p=0.010$) of intratumoral

Table 3. Demographic and histogram apparent diffusion coefficient parameters

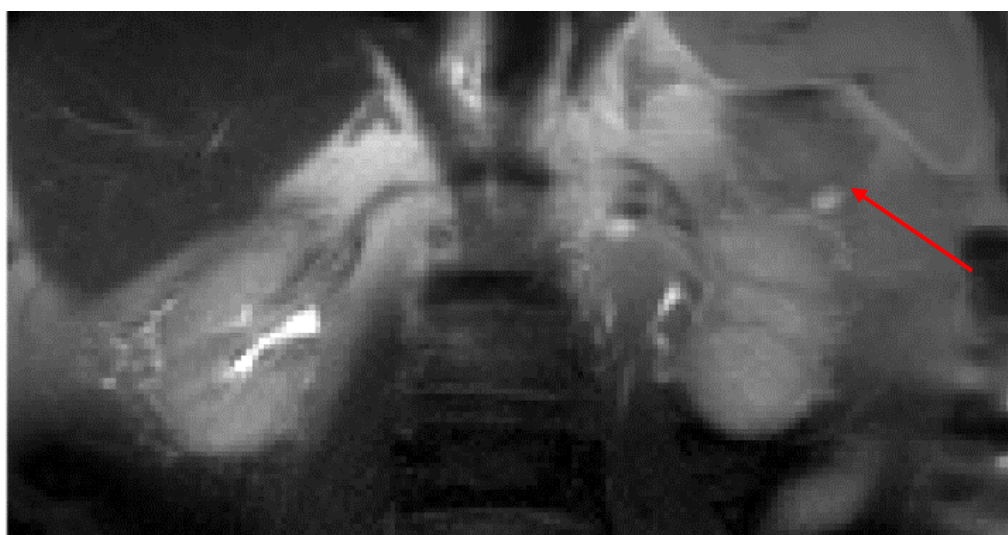
Parameter	Renal Cell Carcinoma	Oncocytoma	<i>p</i> -value*
Sex			0.0012**
Male, No. (%)	27 (84)	3 (43)	
Female, No. (%)	5 (16)	4 (57)	
Age (y)	65 ± 9	65 ± 6	0.294
BMI (kg/m ²)	27.9 ± 3.6	26.9 ± 5.0	0.577
Tumor diameter (mm)	72 ± 33	42 ± 21	0.026
Tumor volume (cc)	242 ± 286	48 ± 59	0.021
<i>ADC</i> histogram parameters:			
Mean	1.51 ± 0.31	1.49 ± 0.59	0.826
Mode	1.35 ± 0.30	1.48 ± 0.59	0.428
Inhomogeneity	0.40 ± 0.14	0.22 ± 0.09	0.006
Skewness	0.37 ± 0.33	0.08 ± 0.42	0.088
Kurtosis	-0.04 ± 0.53	0.07 ± 0.41	0.489
Entropy	3.16 ± 0.16	3.31 ± 0.15	0.027
5 th percentile	0.92 ± 0.29	1.13 ± 0.51	0.218
10 th percentile	1.03 ± 0.27	1.20 ± 0.52	0.258
25 th percentile	1.22 ± 0.27	1.34 ± 0.55	0.389
75 th percentile	1.78 ± 0.39	1.63 ± 0.64	0.826
95 th percentile	2.22 ± 0.48	1.89 ± 0.71	0.231

Unless stated otherwise, values are mean ± SD. Mean, mode, SD and percentiles of apparent diffusion coefficient parameters are in $\times 10^{-3}$ mm²/s. *Group differences were tested with the Mann-Whitney test, unless stated otherwise. **Group difference was tested with the chi-square test.

ADC values were statistically significantly different between OC and RCC. The mean, mode, kurtosis, and the percentiles of the *ADC*-histogram were not significantly different between the two tumor types.(Table 3) No multicollinearity between the significant parameters was present. Therefore, three imaging-derived parameters (tumor volume, and SD and entropy of *ADC* values) were used in the regression model. An ROC-AUC of 0.91 was obtained, providing in a maximum sensitivity and specificity of 86% and 81%, respectively.(Figure 5)

2.3.3 Incremental use of patient demographics

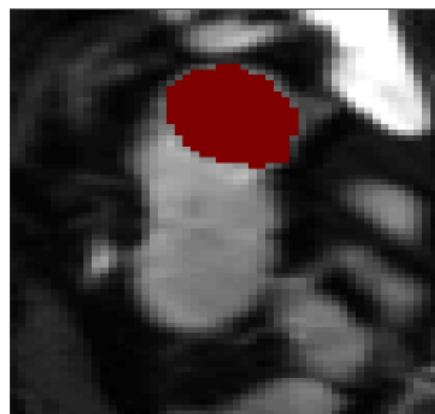
Analysis of demographic parameters showed sex to be statistically significantly different between patients with RCC and OC ($p=0.018$). Patient age and BMI were not significantly different between the two groups.(Table 3) No multicollinearity between the significant *ADC* histogram parameters and sex was observed. The addition of



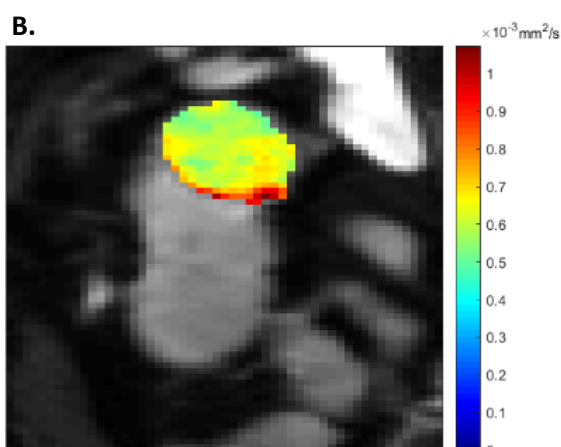
A.



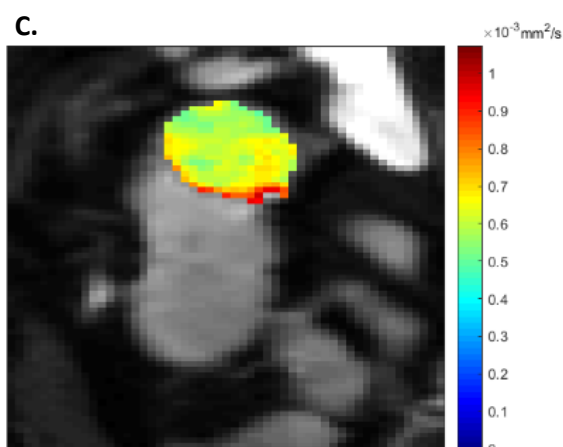
B.



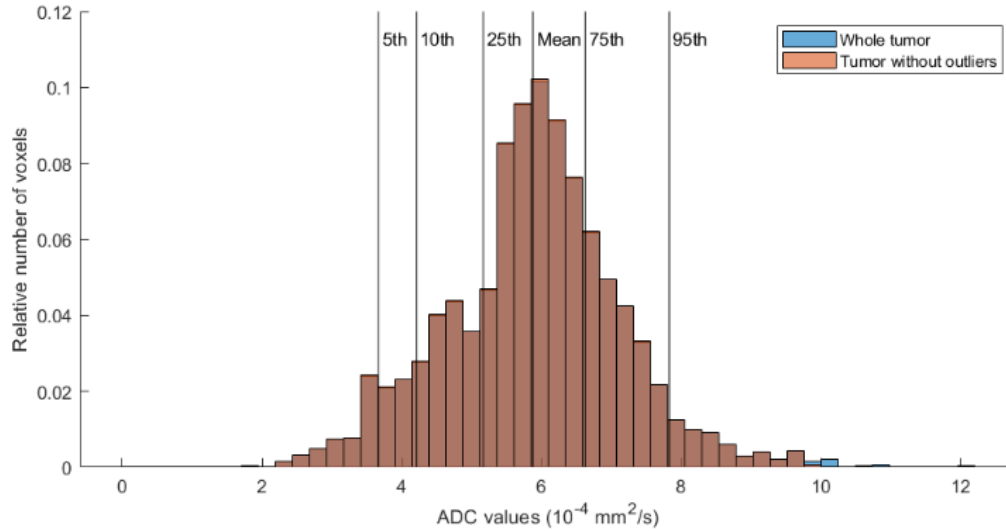
C.



D.



E.



F.

Figure 4. Anatomical coronal T2-weighted image, showing a tumor (oncocytoma, red arrow) in the left kidney.(A). Coronal diffusion weighted image ($b=50 \text{ mm}^2/\text{s}$). (B) Region of interest at one of the slices is shown in red (C). Apparent diffusion coefficient (ADC) map of the tumor (D). ADC map of the tumor after outlier removal (based on volumetric ADC histogram) (E). In (B-E), color bars show ADC values in mm^2/s . In F, the histogram of ADC values in this tumor is shown. ADC values of the whole region of interest and values after removing outliers are shown in blue and red, respectively. Furthermore, the vertical lines represent the mean and percentiles, used in the statistical analysis.

patient sex to the regression model from previous section resulted in ROC-AUC of 0.93.(Figure 5) The maximum sensitivity and specificity were 86% and 84%, respectively.

2.4 Discussion

Preoperative differentiation between benign and malignant renal tumors can prevent high rates of unnecessary surgery. In this work, prediction models combining histogram parameters and patient demographics were evaluated. The histogram analysis revealed statistically significant differences between OC and RCC in SD and entropy of *ADC* values. Li et al. (2018) found statistically significant differences in the entropy of *ADC* values between OC and ccRCC, pRCC and chRCC.[19] They also reported a statistically significant difference in SD between OC and pRCC.

Studies reported that the SD of *ADC* reflects the heterogeneity of a tumor.[37], [38] The heterogeneity originates from angiogenesis, increased cellularity, extravascular extracellular matrix, hemorrhage and necrosis within the tumor. In our study, SD was higher in RCC compared to OC which can imply that malignant tumors are more heterogeneous than OC.

Our results show no significant difference in mean *ADC* between OC and malignant renal tumors. This is in contrast to other studies, who were able to differentiate benign and malignant renal tumors based on the mean *ADC*. [11], [39], [40] Razek et al. used *b*-values 0 and $800 \text{ s}/\text{mm}^2$ and obtained mean *ADC* values of $2.10 \cdot 10^{-3}$ and $1.74 \cdot 10^{-3} \text{ mm}^2/\text{s}$ for OC and solid renal malignancies, respectively.[39] Taouli et al. used *b*-values 0 and $800 \text{ s}/\text{mm}^2$ and obtained mean *ADC* values of $1.91 \cdot 10^{-3}$ and $1.54 \cdot 10^{-3} \text{ mm}^2/\text{s}$ for OC and solid renal malignancies, respectively.[40] While in this study similar results for RCC were found, we obtained considerably lower values for OC. However, the mean *ADC* is susceptible to extreme values which results in a wide range for mean *ADC* values. Furthermore, all of the mentioned studies show an overlap between OC and malignant renal masses. Therefore the mean *ADC* has limited clinical value.

Guidelines state that tumor size is associated with recurrence and development of metastases [3], [41] In this study, there was a significant difference in tumor volume between OC and RCC. Furthermore, we found that RCCs are significantly more common in males, which is in concordance with literature.[21]

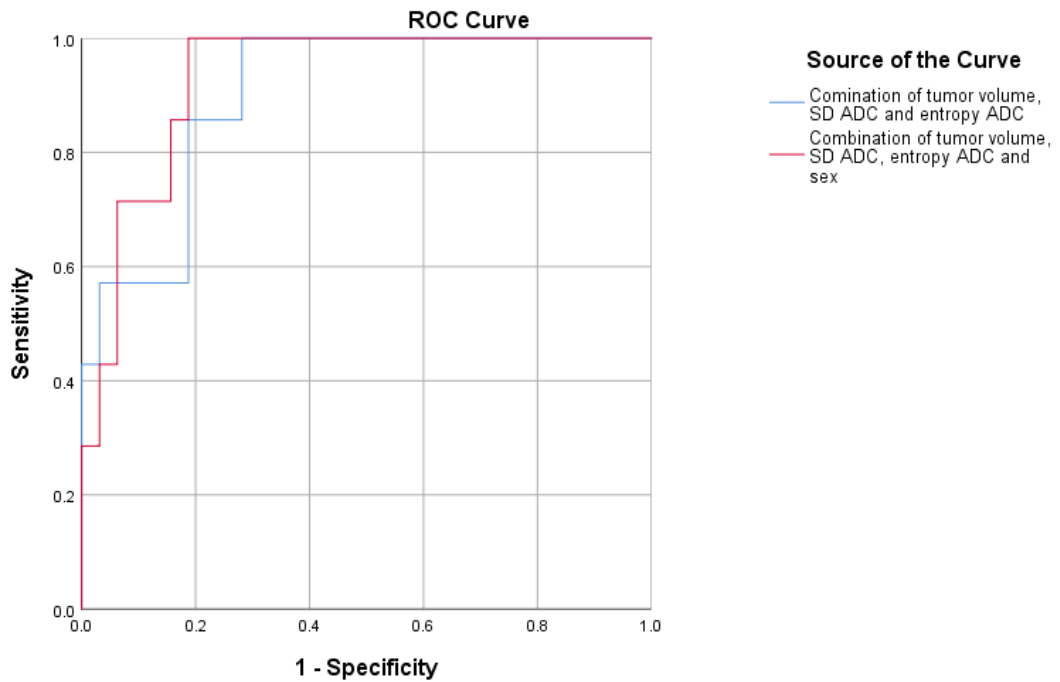


Figure 5. Receiver operator curve of results of both logistic regression models. The result of the combined imaging derived features (tumor volume, standard deviation, and entropy) is shown in blue. The result obtained after addition of the patient demographic parameter (sex) is shown in red.

By combining imaging features (tumor volume, and SD and entropy of *ADC*), we obtained a model that provided a sensitivity of 86% and a specificity of 81% in predicting benign nature of renal tumors. Kim et al. (2016) reported sensitivities of 79.7% and 88.1% and specificities of 44.4 and 33.3% for contrast-enhanced CT and MRI.[24] This indicates that *ADC* histogram analysis provides information that can improve differentiation between OC and RCC. Addition of sex resulted in a small increase in specificity to 84% while maintaining a sensitivity of 86%. This implies that patients' demographic factors are of extra value in this differentiation, although the value is limited.

Our study has several limitations. First of all, we included three patients in which only percutaneous biopsy was used to determine the histologic tumor classification. Percutaneous biopsies are potentially influenced by sampling error and are therefore less reliable compared to histologic analysis after tumor nephrectomy. All of the cases in which biopsy was performed were identified as OC at histopathologic analysis. Due to the small subpopulation, one erroneous classification can have significant consequences in *ADC* value analysis. However, none of these lesions showed significant tumor growth during active surveillance. This is a characteristic that is often found in OC, which increases the likelihood that these tumors are truly OC. Secondly, the study population, and in particular the OC subpopulation, was relatively small. Further studies with larger study populations are required for improving the prediction model. At the moment, the model is likely to be overfit to our data. A larger pool of data is required for model testing and validation. In addition to decreasing the number of unnecessary tumor nephrectomies, an increase in RCC incorrectly being diagnosed as OCC has to be avoided by using the model.

In the current study, the RCC subtypes were grouped. The most prevalent RCC subtypes are clear cell (80%), papillary (10%) and chromophobe (5%).[2] These subtypes all have their own tissue characteristics, which may have caused the broad ranges in parameters within the RCC group in our study. However, our patient population was too small for a separate analysis. Prior studies show promising results in using *ADC* histogram parameters for RCC subtyping.[10], [19], [42], [43] Therefore, research in a larger population has to show whether separate analysis of the subtypes improves RCC subtyping, as well as differentiation of benign and malignant renal tumors.

2.5 Conclusion

The present study evaluated the use of *ADC* histogram parameters and patient demographics for differentiation of solid renal masses. The SD and entropy of *ADC* values, tumor volume and sex are promising parameters in differentiating OC from RCC. The classification model based on these parameters can predict the chance of OC with a sensitivity of 86% and a specificity 84%.

CHAPTER 3 – Evaluation of b -value distributions and fitting algorithms in estimating multi-exponential diffusion-weighted imaging parameters

3.1 Introduction

3.1.1 Diffusion-weighted imaging models

As confirmed by the results of the previous chapter, the use of DWI-imaging shows promising results in the characterization of solid renal masses. However, as in prior studies, reliable differentiation between benign and malignant renal masses is hampered by the wide range in DWI-parameter values, which results in overlapping values between tumor types. The wide range can be attributed to a poor fit of the mono-exponential model (1) to signal intensities acquired with low diffusion weighting. The mono-exponential model assumes that signal decay is only affected by one diffusion component, while randomly oriented perfusion in the microcirculation also induces signal loss when low b -values are used. This randomly oriented perfusion is referred to as pseudodiffusion. To separate diffusion from pseudodiffusion effects, Le Bihan et al. (1986) proposed the intravoxel incoherent motion (IVIM) model [44]:

$$S_b = S_0[(1 - f)e^{-b \cdot D_t} + f \cdot e^{-b \cdot D_p}] \quad (2)$$

This model separates the pseudodiffusion (D_p) from the true tissue diffusion (D_t), while f represents the tissue fraction in which pseudodiffusion takes place. Prior studies have evaluated the use of this model in identifying pathological processes in the kidney.[15], [45]–[49] They showed that IVIM-parameters provide more accurate information about diffusion and are therefore superior over the ADC value in differentiating renal tumors. Mainly f and D_t are considered to be useful for renal tumor discrimination. Though, wide ranges in diffusion parameters still impedes usage of the parameters in clinical decision making.[48] One explanation for wide ranges is that the IVIM model has more degrees of freedom, compared to the mono-exponential model, and therefore robust fitting of the diffusion weighted data can be intricate. More b -values are required for reliable parameter calculation, which elongates the acquisition time. Furthermore, the assumption of Le Bihan et al. comprising the existence of only two sources of signal decay within a voxel has been questioned, because this assumption can impair correct representation of in vivo diffusion processes.[48], [50]

Another multi-exponential model that is increasingly used in characterizing diffusion weighted data is the stretched exponential model. This model is described as follows:

$$S_b = S_0 e^{-(b \cdot DDC)^\alpha} \quad (3)$$

where DDC is the distributed diffusion coefficient, which describes the diffusion rates, and α is the diffusion heterogeneity index, which is related to the degree of intravoxel diffusion heterogeneity. When α is 1, the model is equal to the mono-exponential model, whereas an α close to 0 resembles a high degree of diffusion heterogeneity. The intravoxel inhomogeneity is believed to be higher in tumor types that show high intervoxel heterogeneity.

In contrast to the IVIM-model, the stretched exponential model does not assume a number of distinct diffusion rates. Therefore, there is no direct association between the parameters and biophysical processes. An advantage of the stretched exponential model is that it has only two degrees of freedom, conceivably resulting in a more robust diffusion estimation compared to the IVIM-model.

3.1.2 b -value selection

Low b -values are often used for estimation of DWI-parameters because of the high signal-to-noise ratio (SNR) of these images. Prior studies found that this lead so a dependency between the used b -value distribution and the ADC value, because of the pseudodiffusion contribution.[18], [51] Since multi-exponential models assume more than one source of signal-loss, they are likely to provide a better fit to the data. As a consequence, the dependency between the DWI-parameter and the b -value distribution is expected to be lower in these models. However, fitting these models to the data involves more degrees of freedom in parameter estimation. Therefore, acquisition of more b -value-images is required for reliable parameter estimations.

The effect of pseudodiffusion is largest at low b -values. Since pseudodiffusion is relatively fast compared to diffusion, its effect on the signal intensity and therefore on DWI-parameters is minimal at higher b -values. Thus, to adequately describe the pseudodiffusion effect using the IVIM-model, low b -values have to be acquired extensively. Koh et al. (2011) stated that at least six to eight b -values have to be used in the IVIM-model.[52]

Utilizing more b -values reduces parameter estimation uncertainties and fitting errors. However, this involves longer acquisition times, which adversely affects patient comfort and costs. An optimal b -value distribution would minimize parameter estimation errors and uncertainties, while assuring clinically applicable scanning times.

To date, only two studies have evaluated the use of the stretched exponential model in differentiating renal tumors.[53], [54] Both of these studies were retrospective, and therefore optimal b -value distributions for this model are yet to be discovered.

3.1.3 Data fitting strategies

To acquire DWI-parameters, the multi-exponential models have to be fitted to the signal intensities acquired with different b -values. Previous research regarding DWI in kidneys indicates that fitting algorithms can impact the parameter estimation accuracy in the IVIM model.[55] Methods used to calculate IVIM parameters include segmented fitting techniques and free fitting techniques using least-squares, and Bayesian probability (BP) algorithms.

In the segmented fitting method, the IVIM parameters are estimated subsequently. First, based on the assumption that the signal decay at high b -values is not influenced by pseudodiffusion, D_t is estimated by fitting a mono-exponential function to signal intensities acquired using b -values above a certain cut-off. Hereafter, f is estimated by extrapolating the model to $b = 0$ s/mm², and finally D_p is calculated by fitting the model (with only one degree of freedom left) to all signal intensities. This fitting method is considered to be robust. However, there is no consensus about the cut-off value and estimation of IVIM-parameters is hampered when the estimation results in negative pseudodiffusion fractions.[47], [56] These erroneous values have to be removed, which influences whole tumor DWI-parameter analysis.

Estimation of multi-exponential parameters by (constrained) free fitting is often performed using a maximum-likelihood method with the least-squares criterion. Most commonly, the Levenberg-Marquardt (LM) or trust-region reflective (TRR) algorithms are used. The LM algorithm estimates the IVIM parameters simultaneously by computing a linear approximation of the minimization function in each point. In each iteration, it searches for a point in which the fit error is reduced. The direction and step size are determined using a combination of the Gauss-Newton and steepest descent methods. The LM algorithm does not handle boundary constraints. Therefore, unrealistic parameter estimations can occur. The TRR algorithm is similar to the LM algorithm. However, in TRR, the approximation of the loss function is only trusted in a region near the current iteration. At each iteration, the size of the trust region is determined by how well the loss model fits to the original problem. When the fit results in a satisfactory error reduction, the model is a good representation of the original problem. In this case the trust region for the next iteration is enlarged. Only after defining the step size, the step direction is determined. In this way, TRR algorithms are more likely to converge to the global minimum. Furthermore, boundary constraints can be incorporated easily. For these reasons, the TRR algorithm is preferred over the LM algorithm.

Both, LM and TRR algorithms assume that the parameters of interest are unknown but unique: there is one true parameter. The BP algorithm assumes that the parameters to be estimated (φ), as well as the data (D), have a probability distribution: $P(\varphi|D)$. To determine the probability of φ , prior knowledge about the distribution of the parameters, $P(\varphi)$, is required. This is called the prior distribution. Most studies use noninformative or low-informative priors (priors with a large variance) to prevent them from influencing the posterior probability excessively.[57] To estimate the posterior distribution, the prior estimation is updated using the measured data on the basis of approximate methods like the Markov chain Monte Carlo method. The essence of the BP algorithm is to maximize the posterior probability.[58] To describe the results of the Bayesian fit, marginal probability distributions are summarized in terms of central tendency. Prior studies used the mean or mode of the parameter distributions.[57], [59], [60]

3.1.4 Diffusion weighted magnetic resonance imaging sequence

There are three major challenges in acquiring DW-images: the intrinsically low SNR, acquisition time and data fidelity. Most commonly, an SS-EPI sequence is used for acquiring DW-images. To accelerate this sequence, parallel imaging techniques are often used. Most of these techniques, including sensitivity encoding (SENSE) and generalized auto-calibrating partially parallel acquisitions (GRAPPA), acquire less phase-encoding lines which leads to aliasing effects. These effects are then unfolded during the reconstruction process by using the encoding power of the radiofrequency receiver coil arrays. The geometry factor (g -factor) is a measure representing how well the

signals can be unfolded. It expresses the spatially varying noise amplification. The acceleration factor (R , number of lines skipped in k-space) and g -factor affect the SNR of the images according to (4).[61]

$$SNR_{parallel} = \frac{SNR}{g\sqrt{R}} \quad (4)$$

In addition to the decrease in SNR, using an acceleration factor of 2 only results in a reduction in acquisition time of 20-30% due to the relative long phase preparation and readout durations.[62] By exciting multiple slices simultaneously, simultaneous multislice (SMS) techniques provide true acceleration and mitigate the SNR loss. As a result of the acceleration along the slice direction, the acceleration factor is determined by the size of the excited volume.[63] The main limitation is the ability to properly unfold the aliased slices due to the high demand on spatial variations in coil sensitivities, resulting in high g -factor noise. This can be improved by using the Controlled Aliasing In Parallel Imaging Results IN Higher Acceleration (CAIPIRINHA) technique, in which the phase of the individual radiofrequency pulses is alternated. Therefore, variations in coil sensitivity are utilized in both, in- and through-plane directions.[63]

3.1.5 Research aim

Precise and robust calculation of diffusion parameters remains challenging. Using an optimal set of b -values combined with a fitting algorithm that provides robust estimation of the parameters is believed to improve precision and accuracy of the DWI parameters. Therefore, the first two aims of this study were to optimize b -value sampling for multi-exponential analysis of DWI data and to compare fitting algorithm performances. To verify the results, the acquired protocol was tested by acquiring DWI MRI in healthy volunteers. Furthermore, since acquiring DWI images requires long acquisition times, the performance of simultaneous multislice imaging will be compared to the standard acquisition protocol in terms of acquisition time, image quality (SNR) and parameter estimation.

3.2 Methods

3.2.1 Simulations

All computer simulations to optimize the DWI protocol were implemented in MATLAB (2019a, Mathworks Inc, Natick). First, initial b -value distributions containing three b -values were obtained for each combination of DWI-model and fitting algorithm. All sets of three unique b -values were tested, using b -values in the range of 0 to 800 in steps of 30 s/mm² and 900 to 1400 in steps of 100 s/mm². The following diffusion parameters were used: $S_0 = 900$, $f = 30\%$, $D_p = 50 \cdot 10^{-3}$ mm²/s, $D_t = 1.2 \cdot 10^{-3}$ mm²/s, $DDC = 2.1 \cdot 10^{-3}$ mm²/s, and $\alpha = 0.67$. [18], [64], [65] For each b -value set, signal values were generated according to (1), (2) and (3). The simulated signals were Rician distributed using noise with a standard deviation (SD) of $0.01 \cdot S_0$ as described by Zhang et al. (2016).[66]

Subsequently, the DWI-models were fitted to the signal intensities using the TRR and BP algorithm. Initial parameter estimates were: $f = 10\%$, $D_p = 10 \cdot 10^{-3}$ mm²/s, $D_t = 1 \cdot 10^{-3}$ mm²/s, $DDC = 1.0 \cdot 10^{-3}$ mm²/s, and $\alpha = 0.8$. For each set of three b -values, the calculation and fitting of the signal was repeated $N = 1000$ times.

For each DWI-model, the set of b -values with the least overall error was now used as the starting b -value distribution. The initial b -value set was extended with one additional b -value in the range of 0 to 800 s/mm² in steps of 10 s/mm² and in the range of 850 to 1400 s/mm² in steps of 50 s/mm². Again, the overall error was calculated using 1000 repetitions per b -value set. The set with the smallest overall error became the new initial b -value set. This process was repeated 22 times to obtain a b -value set of 25 b -values for each DWI-model. In

Table 4. Magnetic resonance imaging parameters

Parameter	Axial T2-weighted HASTE	Coronal T2-weighted Turbo Spin Echo	Clinical DWI	Accelerated DWI
No. of slices	30	35	35	35
TR (ms)	2000	2750	2300	2100
TE (ms)	93	104	48	51
Flip angle (°)	150	140	90	90
FOV (mm x mm)	450 x 450	400 x 400	308 x 380	308 x 380
Voxel size (mm)	1.4 x 1.4 x 4	1.3 x 1.3 x 5	2 x 2 x 4	1.5 x 1.5 x 5
Bandwidth (Hz/Px)	781	260	1735	2440
b -values (s/mm ²)	NA	NA	0, 60, 90, 110, 150, 400, 800, 1400	0, 30, 60, 110, 150, 400, 800, 1040

TR: repetition time, TE: echo time, NA: not applicable

Bayesian fitting, the estimated parameters are expressed using their posterior probability. In the current study, the mean and mode were used to describe the central tendency.

To evaluate the performance of the multi-exponential models, all simulations were repeated using the mono-exponential model. An ADC -value of $1.59 \cdot 10^{-3} \text{ s/mm}^2$ was used for generating the DWI-signal intensities and an ADC of $1.0 \cdot 10^{-3} \text{ mm}^2/\text{s}$ was used as initial parameter estimate. The mono-exponential model was fitted to the data using the TRR algorithm.

A summary of design parameters for the simulations can be found in Appendix A.

3.2.2 Algorithm optimization criteria

The performance of the fitting algorithms was assessed with respect to adequacy in parameter estimation and robustness. The relative fitting error (ε_φ) between the estimated parameter at the i^{th} repetition (φ_i) and the ground truth parameter (φ_{true}) used for generating S_b was calculated using:

$$\varepsilon_\varphi = \frac{\sqrt{\frac{1}{N} \sum_{i=1}^N (\varphi_i - \varphi_{true})^2}}{\varphi_{true}} \quad (5)$$

The model parameter estimation error was defined as the sum of the individual relative errors:

$$\varepsilon_{total,model} = \sum \varepsilon_\varphi \quad (6)$$

The robustness of fitting was evaluated by calculating the SD of each parameter over the N iterations. An optimal b -value set results in a low parameter estimation error combined with a low SD. Therefore, a weighted sum was used as a measure for overall model error per b -value distribution. This was defined as:

$$E_{b,model} = \frac{\varepsilon_{total,model}}{\min(\varepsilon_{total,model})} + \frac{SD_{model}}{\min(SD_{model})} \quad (7)$$

3.2.3 Combined b -value distribution selection

The optimized b -value set for the combination of the IVIM and stretched exponential model was evaluated for both algorithms. For this, the b -values that belong to the individual optimized b -value distributions were evaluated in another simulation. In this simulation, all combinations of b -values from the optimized b -value distributions were tested. Signal values were generated in the same way as in the former simulation and both DWI-models were fitted to the data using the TRR and BP algorithms. Fitting errors of both algorithms were joined to one error measure for the combined models according to (8). The set of b -values resulting in the smallest combined error was used as overall optimized b -value distribution.

$$E_{b,combined} = E_{b,IVIM} + E_{b,stretched \text{ exponential}} \quad (8)$$

3.2.4 Human subject validation

To test the results obtained by the simulations, abdominal DWIs were acquired in four healthy volunteers (mean age, 24 years; ranging from 24 to 25). All scans were performed with a 3T MRI scanner (Magnetom prisma, Siemens Healthcare, Erlangen, Germany) and an 18-channel body and a spine phased-array coil. The scanning protocol consisted of the following sequences: an axial T2-weighted fat-suppressed HASTE in breath-hold and a coronal T2-weighted fat-suppressed turbo spin echo in expiration. In all four volunteers, the standard renal DWI sequence was used with additional b -values. This sequence consists of a coronal, respiratory-triggered SS-EPI with GRAPPA acceleration factor 3. b -values of 0, 60, 90, 110, 140, 400, 800, and 1400 s/mm^2 were used. b_0 -images were acquired with 1 repetition, while all other images were acquired with 2 repetitions. Further scan parameters are shown in Table 4.

In two of the subjects, additional, accelerated DWI was performed. This sequence consisted of a respiratory-triggered SS-EPI sequence using blipped-CAIPIRINHA SMS imaging with an acceleration factor of 2, combined with a GRAPPA factor of 2 for in-plane acceleration (Siemens WIP prototype sequence no. 996C). During b -value image reconstruction of this DWI-sequence, motion correction was performed. This includes applying the following series of correction algorithms: in-plane registration, filtering, denoising and rescaling to compensate for signal loss due to motion. The same b -values were used as in the first DWI sequence except for $b = 90 \text{ s/mm}^2$, which was replaced for $b = 30 \text{ s/mm}^2$, and $b = 1400 \text{ s/mm}^2$, which was replaced for $b = 1040 \text{ s/mm}^2$. These adjustments were

performed, since signal intensities acquired with b_{90} deviated in the signal decay curve and b_{1040} was the highest b -value that could be used in this sequence. b_0 images were acquired using 1 repetition, b_{30} to b_{400} were acquired with 2 repetitions, and higher b -value images were acquired using 4 repetitions. For both respiratory triggered sequences, a liver dome scout was used and diffusion gradients were applied in three directions.

To estimate DWI-parameters, ROIs were drawn manually covering the cortex of both kidneys in 3D, using FireVoxel. The T2-weighted images were used as anatomical landmark for this segmentation. In MATLAB, the TRR and BP algorithms were used for voxel-wise fitting of the DWI models to the signal intensities in the ROI. Start values for the DWI-models were the same as during simulations. Mean parameter values were calculated and compared between the algorithms. Furthermore, the SNRs of $b = 0$ and 800 s/mm^2 images, obtained with both DWI-sequences, were compared. The SNR was calculated using:

$$SNR = \frac{S_0}{\sigma_{bg}} \quad (9)$$

where σ_{bg} is the standard deviation of signal values in the background. S_0 and σ_{bg} were derived within the same ROI for both DWI sequences.

3.3 Results

3.3.1 Simulations

In Table 5, optimized b -value distributions are listed for each of the diffusion weighted models and fitting algorithms. The values are listed in order of addition to the b -value distribution. The first three b -values were the initial optimal b -value set. These initial b -value sets contain two low and one high b -value for the BP algorithm, while this was one low and two high b -values for the TRR algorithm in both the IVIM and the stretched exponential model. The histograms of the b -value distributions of the models are shown in Figure 6. A higher frequency indicates that more b -values in that range were used to calculate the diffusion weighted parameters. In both multi-exponential models, over 50% of the b -values are below 200 s/mm^2 , regardless of the algorithm that was used to calculate the parameters. Additionally, a few medium-high and high b -values ($b > 200 \text{ s/mm}^2$) were used for model fitting.

Parameter fitting errors and standard deviations per number of b -values added to the initial b -value sets are shown in Figure 7, together with their gradients. The BP algorithm, provided a more adequate fit for the IVIM model compared to the TRR algorithm. Parameter estimation errors were smallest when using the mode as central tendency measure. In the stretched exponential model, there was no clear difference between the algorithms. When adding five b -values to the initial set of three b -values, both algorithms show an inflection point in estimation error decline. Therefore, a b -value distribution with a total of eight b -values was used in human subject validation. Fitting errors and standard deviations per DWI-parameter can be found in Appendix B. Parameter errors and standard deviations for the optimized b -value distributions are provided in Table C1.

Table 6 lists optimized b -value distributions for the combinations of the IVIM and stretched exponential models. The smallest parameter estimation error was obtained when the BP algorithm with mean as central tendency measure was used for both models. Parameter errors and standard deviations for all combinations of algorithms and DWI models are listed in Table C2.

3.3.2 In vivo measurements

None of the scans had severe artifacts that hampered diffusion parameter estimations for the renal cortex. Therefore, DWI parameters were calculated in eight kidneys for the standard DWI-sequence and in two for the accelerated sequence. The ROIs had a mean size of 23.3 cm^3 (range: $17.3 - 28.0 \text{ cm}^3$).

Table 7 shows mean values, SD and ranges of diffusion parameters. Considering the multi-exponential DWI-parameters that describe diffusion, D_t shows higher means than DDC . Comparing these diffusion parameters with the ADC value, DDC is closer to the ADC .

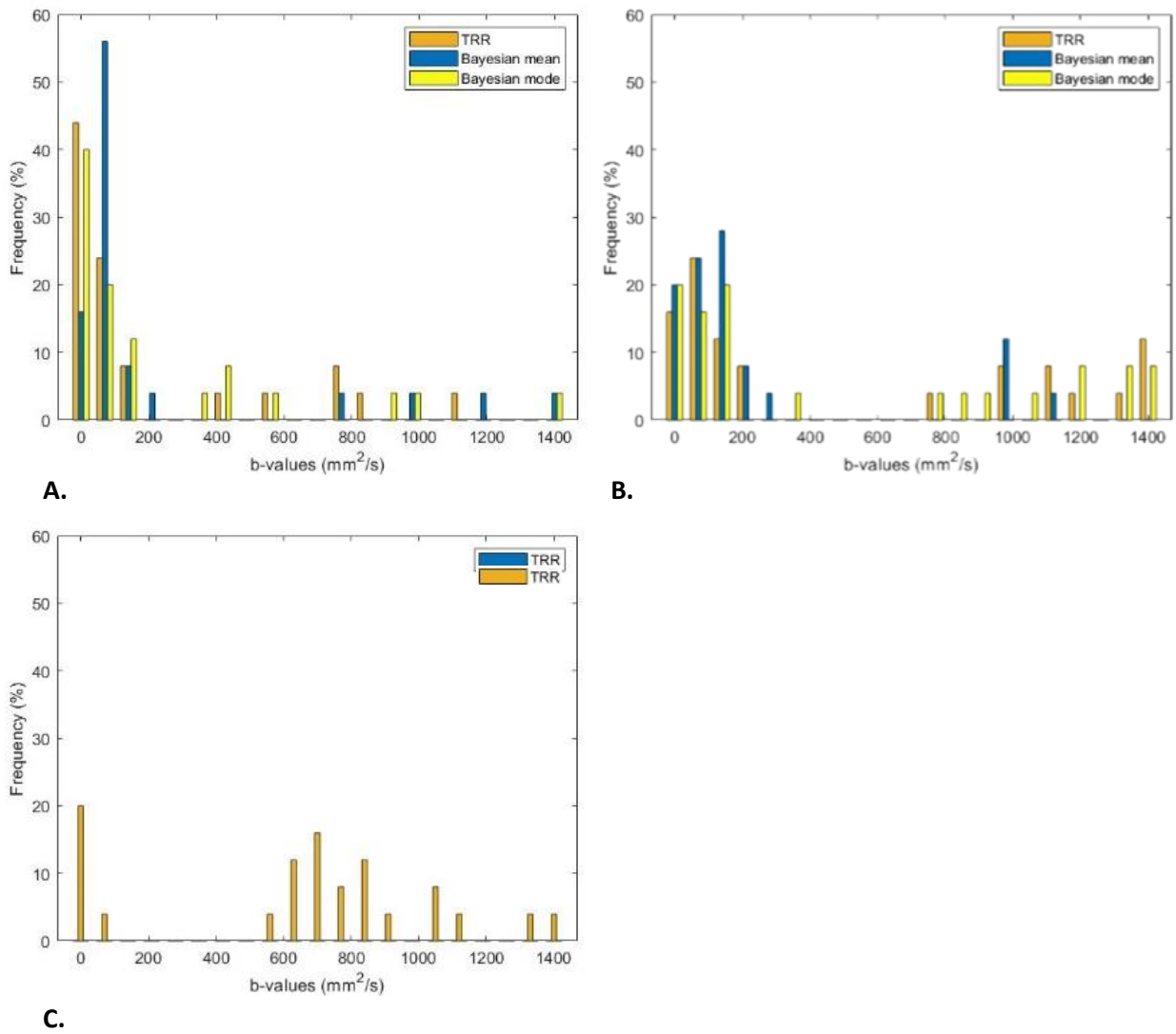


Figure 6. Histograms showing the frequency of b -values used in model fitting, for the intravoxel incoherent motion (A), stretched exponential (B), and mono-exponential model (C).

Table 5. b -value distributions for all DWI-models and fitting algorithms

	Initial b -value set (s/mm^2)	Added b -value sets (s/mm^2)
Intravoxel incoherent motion		
TRR	0, 900, 1100	430, 10, 110, 20, 110, 10, 60, 0, 160, 20, 60, 130, 130, 790, 30, 130, 0, 20, 180, 850, 590, 80
BP mean	0, 90, 1200	110, 140, 30, 800, 90, 110, 100, 90, 1200, 130, 100, 90, 100, 120, 130, 140, 1400, 50, 100, 130, 10, 1000, 1400
BP mode	0, 90, 1400	10, 120, 10, 140, 170, 30, 20, 420, 100, 1000, 0, 50, 900, 130, 90, 400, 30, 30, 140, 570, 20, 440
Stretched exponential		
TRR	0, 1000, 1300	110, 140, 1350, 160, 230, 1100, 0, 1350, 120, 0, 110, 1000, 1200, 210, 0, 130, 1100, 1250, 80, 160, 750, 180
BP mean	0, 110, 1000	60, 110, 140, 120, 150, 1000, 90, 30, 190, 150, 1100, 80, 230, 110, 140, 50, 170, 1000, 320, 240, 10, 170
BP mode	0, 90, 1400	90, 1300, 170, 1150, 0, 190, 100, 900, 0, 150, 1300, 120, 770, 180, 0, 850, 0, 340, 150, 1200, 1050, 1350
Mono-exponential		
TRR	480, 540, 630	10, 750, 580, 0, 760, 730, 790, 1300, 40, 700, 750, 670, 620, 1050, 1200, 50, 950, 950, 850, 630, 30, 120

TRR: trust region reflective algorithm, BP: Bayesian probability algorithm

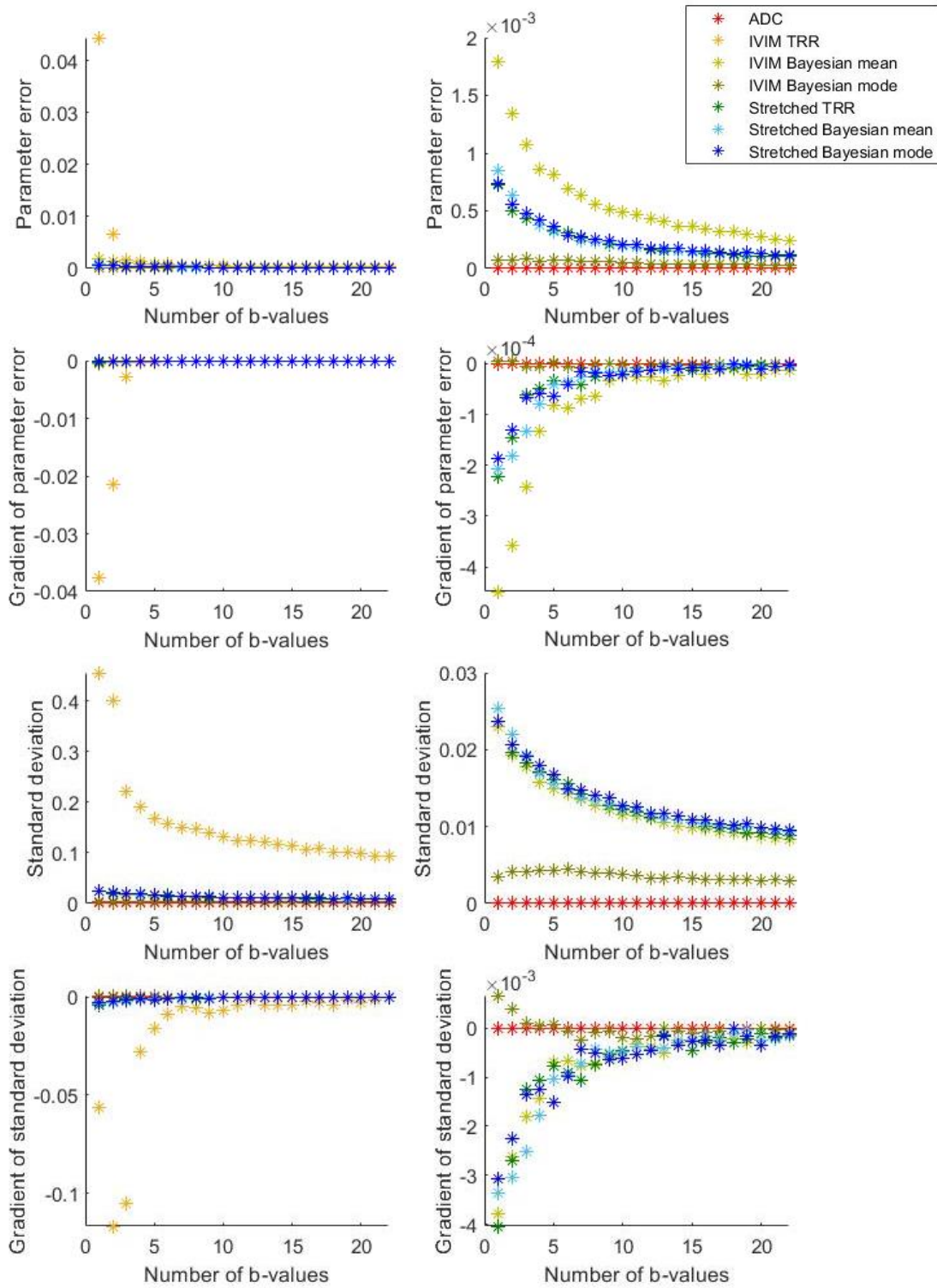


Figure 7. Total fitting errors and standard deviations per number of added b -values, for each of the algorithms. Top row: errors in parameter fitting, Second row: Gradient of fitting errors, Third row: Standard deviations of fitted parameters, Bottom row: Gradient of standard deviations. The graphs in the right column show data obtained with five or more b -values, and data of the TRR algorithm for the IVIM model are removed because of their large errors. ADC: apparent diffusion coefficient, IVIM: intravoxel incoherent motion, TRR: trust region reflective algorithm.

For the IVIM model, all parameter estimates depend on the fitting algorithm that was used. Compared to the BP algorithm, the TRR algorithm resulted in larger standard deviations for all parameters considering the standard DWI sequence, and for D_p considering the accelerated DWI-sequence. For the stretched exponential model, there was no clear dependence between the mean and standard deviations and the fitting algorithm that was used.

In Figure 8 parameter maps are shown for all DWI-parameters, obtained using the BP algorithm. Parameter maps obtained with the accelerated sequence are smoother, which can be predominantly seen in f and D_p maps. This results in a better anatomical delineation of the renal pelvis.

Regardless of the fitting algorithm that is used, the mean of f is lower, and the mean of D_t is higher in the accelerated sequence compared to the standard sequence. The differences in D_p between both sequences, depend on the fitting algorithm that was used. Means and standard deviations of stretched exponential DWI parameters were similar for both DWI-sequences.

For both sequences, the SNRs calculated in the dorsal part of the kidney were higher compared to those calculated in the central cross section of the kidney in the coronal plane. The SNRs in dorsal and central coronal cross-sections of $b = 0$ s/mm² images were 135 and 53 for the standard sequence, compared to 356 and 119 for the accelerated sequence. In $b = 800$ s/mm² images, the SNRs were 29 and 25 for the standard sequence, and 46 and 44 for the accelerated sequence, respectively. The mean acquisition times were 15:55 min. (range: 14:39 to 18:09 min.) and 1:44 (1:54 and 1:33 min.) for the standard and accelerated sequences, respectively.

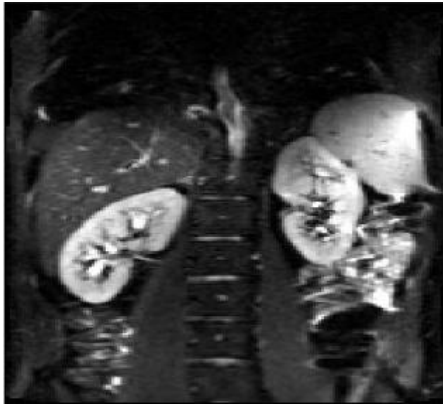
Table 6. Optimized b -value distributions for the combination of the IVIM and stretched exponential models

Model combination		b -value distribution
IVIM	Stretched	
BP mean	TRR	0, 90, 190, 230, 800, 1000, 1200, 1400
BP mean	BP mean	0, 60, 90, 110, 140, 400, 800, 1400
BP mean	BP mode	0, 30, 60, 90, 140, 800, 1000, 1200, 1400
BP mode	TRR	0, 30, 90, 110, 800, 1000, 1200, 1400
BP mode	BP mean	0, 90, 190, 230, 800, 100, 1200, 1400
BP mode	BP mode	0, 60, 110, 140, 230, 1000, 1200, 1400

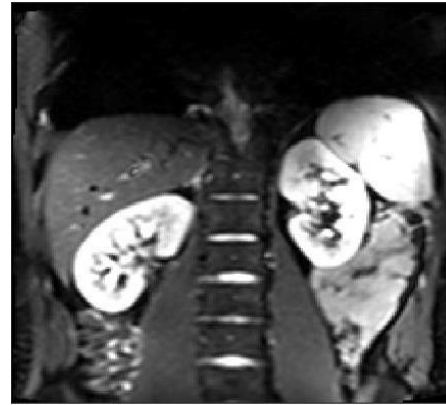
BP: Bayesian probability, TRR: trust-region reflective

Table 7. Mean Diffusion parameters in renal cortex of healthy volunteers

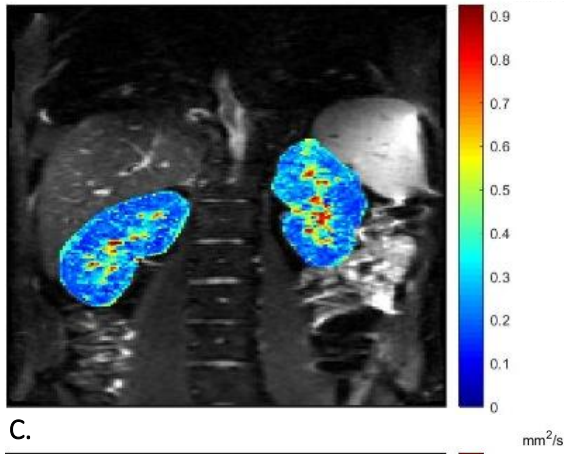
	Standard (n=4) ROI mean	SMS (n=2) ROI mean
Volume ROI (cm³)	23.7 ± 3.3 (17.3-28.0)	22.5 ± 2.3 (19.2-24.1)
f (%)		
TRR	39.8 ± 7.8 (32.8-53.4)	22.7 ± 6.0 (16.4-28.5)
BP mean	26.6 ± 2.8 (22.0-29.3)	21.8 ± 5.1 (17.2-26.4)
BP mode	20.7 ± 5.4 (10.8-25.9)	15.7 ± 7.9 (7.7-22.6)
D_p ($\cdot 10^{-3}$ mm²/s)		
TRR	109 ± 41.9 (41.3-165.7)	92.1 ± 29.4 (66.5-124.8)
BP mean	29.7 ± 2.4 (26.0-34.4)	32.8 ± 4.9 (27.2-37.0)
BP mode	10.0 ± 3.0 (5.5-13.2)	12.3 ± 2.8 (9.5-16.0)
D_t ($\cdot 10^{-3}$ mm²/s)		
TRR	1.34 ± 0.23 (0.94-1.51)	1.69 ± 0.06 (1.61-1.74)
BP mean	1.68 ± 0.06 (1.58-1.74)	1.75 ± 0.05 (1.71-1.82)
BP mode	1.71 ± 0.06 (1.61-1.80)	1.73 ± 0.08 (1.62-1.83)
DDC ($\cdot 10^{-3}$ mm²/s)		
TRR	2.33 ± 0.19 (1.99-2.52)	2.38 ± 0.24 (2.14-2.68)
BP mean	2.34 ± 0.19 (2.02-2.53)	2.39 ± 0.23 (2.16-2.68)
BP mode	2.31 ± 0.24 (1.98-2.51)	2.36 ± 0.23 (2.13-2.65)
α		
TRR	0.80 ± 0.05 (0.76-0.87)	0.77 ± 0.05 (0.72-0.82)
BP mean	0.79 ± 0.04 (0.76-0.87)	0.76 ± 0.04 (0.72-0.81)
BP mode	0.80 ± 0.04 (0.76-0.85)	0.76 ± 0.05 (0.72-0.81)
ADC ($\cdot 10^{-3}$ mm²/s)		
TRR	2.14 ± 0.14 (1.88-2.29)	2.18 ± 0.19 (2.02-2.41)



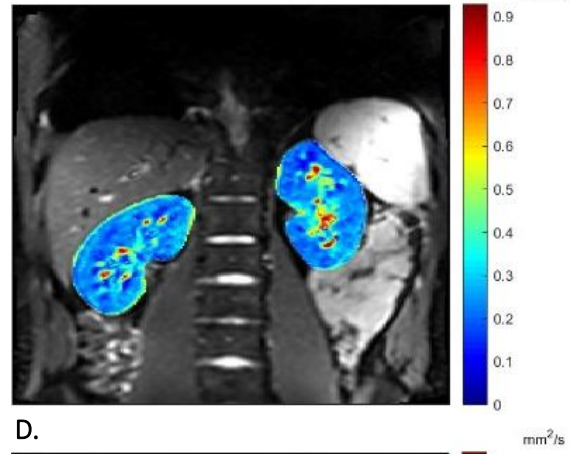
A.



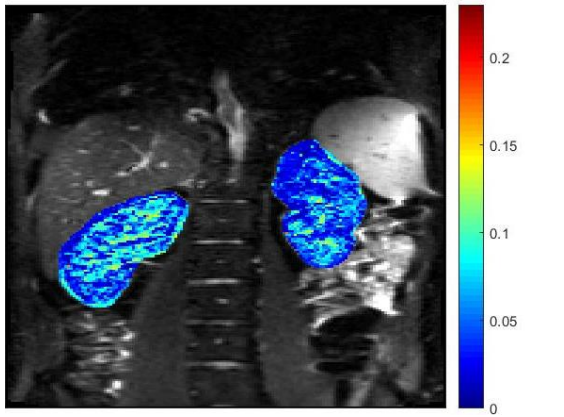
B.



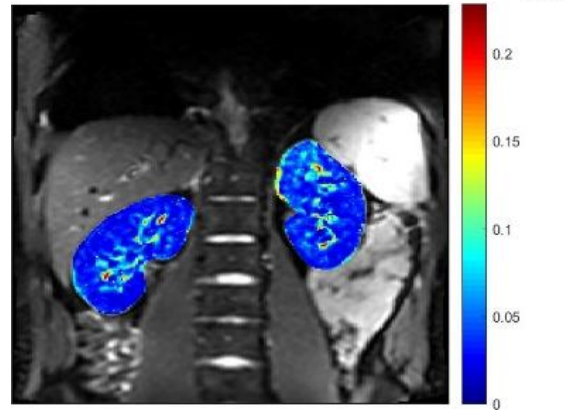
C.



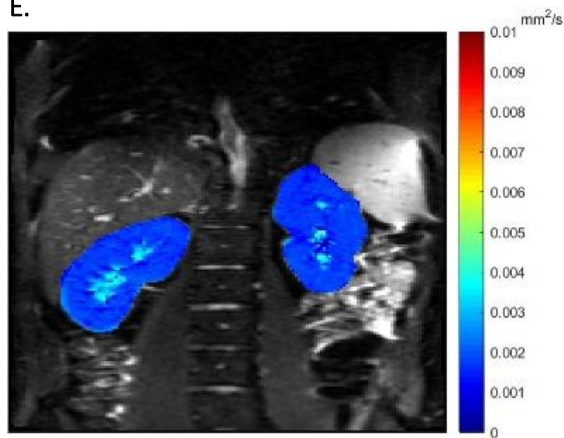
D.



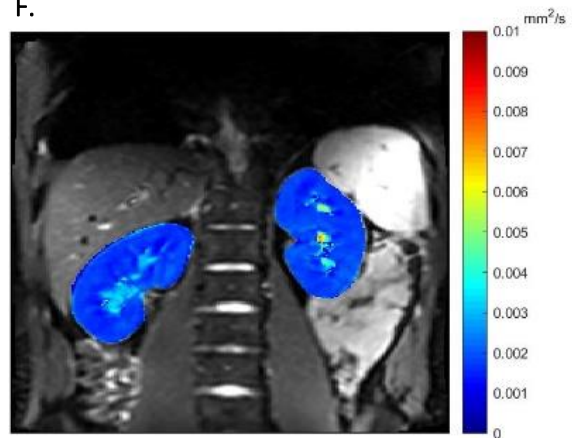
E.



F.



G.



H.

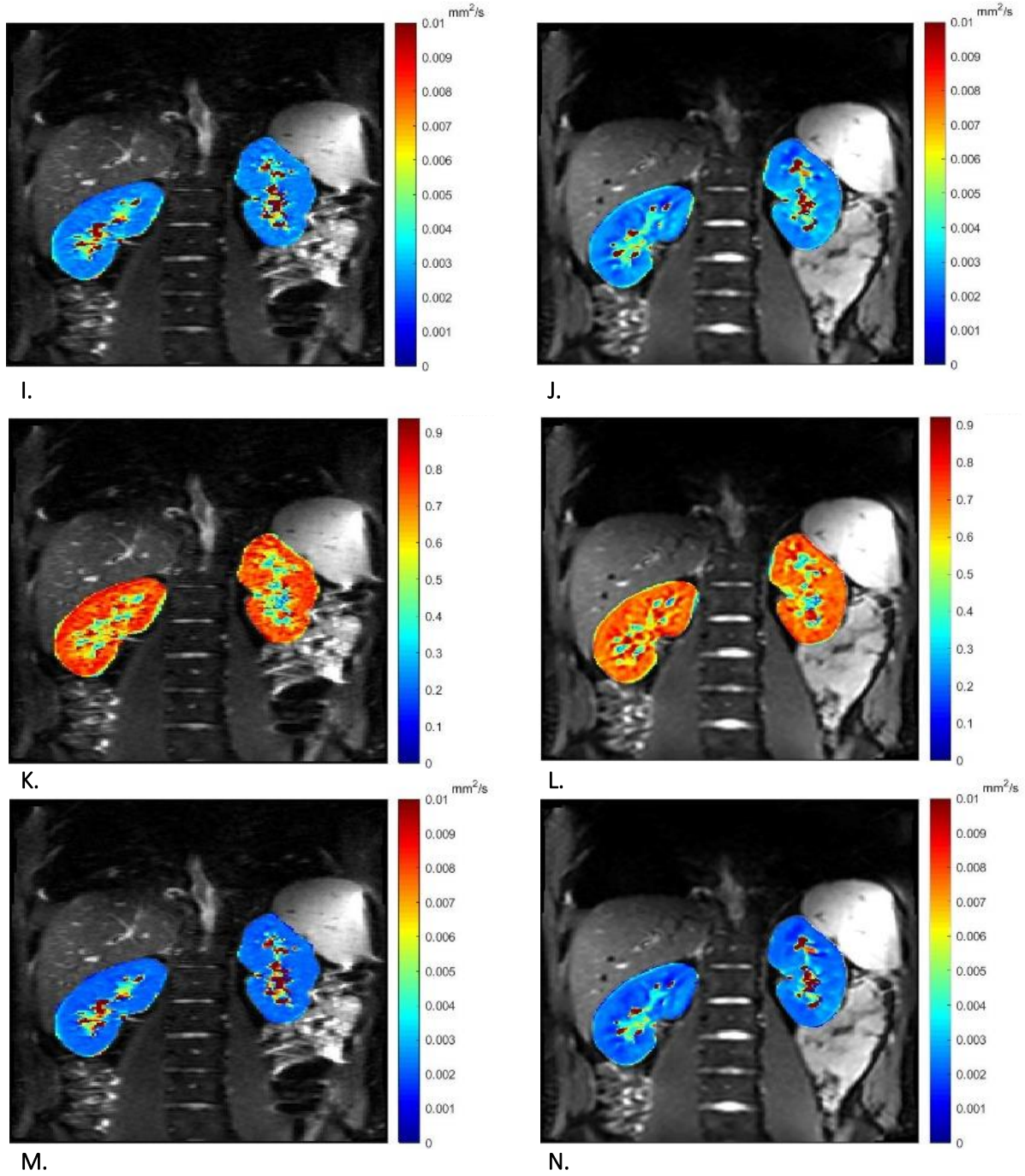


Figure 8. Comparisons of renal multi-exponential DWI-parameter maps acquired using the standard (left column) and accelerated (right column) DW MRI-sequences. Voxel-wise model fitting was performed using the Bayesian probability algorithm for multi-exponential DWI-model parameters, with mean as central tendency measure. (A) and (B) show the b_0 images of both sequences. Parameter maps for IVIM-parameters f , D_p , and D_t are shown in (C/D), (E/F), and (G/H), respectively. Parameter maps for stretched exponential parameters DDC and α are shown in (I/J) and (K/L), respectively. In (M/N), apparent diffusion coefficient parameter maps are shown.

3.4 Discussion

3.4.1 Simulations

In this study optimal b -value distributions were examined for two DWI models and two fitting algorithms. A b -value distribution containing eight b -values was found to be optimal. The fitting algorithm was found to have an effect on parameter estimations in the IVIM model, with the BP algorithm providing more robust fitting compared to the TRR algorithm. The used fitting algorithm had less influence on parameter estimation in the stretched exponential model. Results from in vivo data were in agreement with the results from the simulations.

The optimal b -value distributions were acquired for the IVIM and stretched exponential models using Monte-Carlo simulations. In the optimized b -value distributions, low b -values are sampled extensively compared to high b -values. A proper fit was obtained with the use of only one or two high b -values (depending on the fitting algorithm). Using these distributions the bias in parameter estimations is small, small which means that the accuracy of the algorithms in parameter estimation is high. The standard deviations, which reflect a measure for precision, were relatively large. Therefore, the ranges in parameter values can be large. However, absolute values for standard deviations are small. When considering the parameters that describe tissue diffusion (D_t and DDC), errors in estimation are comparable when the BP fitting algorithm is used. These errors are also comparable to the estimation error for the ADC value of the mono-exponential model.

The error in parameter estimation decreases with increasing numbers of sampled b -values. Using many b -value images for DWI-parameter estimations, requires long acquisition times and would therefore adversely affect patient comfort and costs. Consequently, a trade-off has to be made between the gain in parameter estimation accuracy by adding an extra b -value and the increase in acquisition time. In the current study, this trade-off was made based on the inflection point in the parameter estimation error graphs. For the IVIM-model, as well as the stretched exponential model, first inflection points were found when five b -values were added to the initial b -value sets, resulting in optimized b -value distributions containing eight b -values. This finding is in accordance to the study of Koh et al. (2011) who stated that six to eight b -values should be used with several signal averages to properly describe the signal decay using multi-exponential DWI-models.[52]

When comparing errors in IVIM parameter estimation between the TRR and the BP algorithm, the biases are approximately a factor 10 smaller when the BP algorithm was used. However, the variability in the parameter estimates is much higher, especially the variability in D_t , which is more than a factor 1000 larger when the TRR algorithm was used. These results are in concordance with the study of While (2017), who compared Bayesian and LSQ fitting in simulations for liver and breast tissue.[67] The high variability can result in wide parameter ranges in in vivo measurements. Therefore, the BP algorithm is more suitable for IVIM parameter estimations.

In the BP algorithm, the biases are approximately a factor 10 smaller when the mode is used as central tendency measure, compared to the mean when considering the optimized b -value distribution for all IVIM parameters separately. This can be explained by the fact that the posterior probabilities of the parameters are skewed. Gustafsson et al. (2018) described that using the mode as a measure for central tendency resulted in reduces biases.[57] However, they also found that the variability was higher when the mode was used compared to the mean. The latter was not found in the performed simulations, since the variability was a factor three smaller when the mode was used. However, absolute differences in bias and variability between the central tendency measures are small.

In the stretched exponential model, the TRR and BP algorithms provided parameter estimates with comparable bias and variability. These DWI-parameters therefore seem to be less dependent on the fitting algorithm that is used.

3.4.2 In vivo validation

The errors in parameter estimations, obtained using the b -value distribution optimized for the combination of the IVIM and stretched exponential model, were similar to the errors obtained with the model specific optimized b -value distributions. Therefore, this b -value distribution was used for DWI-acquisition in healthy volunteers.

The findings obtained in in vivo measurements are consistent with the results of the simulations. In the IVIM model, the parameter estimates obtained with the TRR algorithm are substantially different from those obtained with the BP algorithm. Standard deviations, and thus variation in parameters, are larger when the TRR algorithm

is used. Therefore, IVIM parameter estimation using the TRR algorithm is less suitable for describing tissue diffusion and ultimately differentiation of renal tumors. Prior studies have frequently used the combination of the DWI-model and fitting algorithm. However, most studies use tighter parameter constraints during fitting or remove parameter values from analysis that are believed to be outside ‘realistic’ parameter range for renal DWI after data fitting.[20], [68]–[70] These methods provide more robust fitting results, but might also obscure true tissue diffusion effects since in vivo diffusion parameters are unknown.

The in vivo IVIM-parameters are in concordance with results of prior studies concerning healthy renal cortex.[71], [72]. Zhang et al. (2009) obtained slightly smaller variances in f , while variances in D_p are similar and variances in D_t were larger compared to our results.[71] However, they used 27 b -values during parameter fitting, which is expected to provide better fits. The parameter variations in the study of Notohamiprodjo et al. (2015) were similar for f , and larger for D_p and D_t compared to our results.[72] They used 10 b -values, acquired along 20 diffusion directions and calculated the parameters using FireVoxel.

In accordance with the results of the simulations, the TRR and BP algorithms produce similar parameter estimation results for the stretched exponential model. Based on simulation and in vivo results, the stretched exponential model is able to provide robust DWI-parameter estimations for renal cortex. To the best of our knowledge, stretched exponential DWI parameters in healthy kidney tissue are unknown. Although DDC values are higher compared to D_t , they have the same order of magnitude and could therefore be representative for tissue diffusion.

In general, standard deviations of in vivo DWI-parameters are considerably larger compared to the those obtained with simulations. This was expected since in vivo signal decay most likely does not follow model equations exactly, contrarily to the simulations. It is likely that there are physiological differences in the degree of diffusion taking place between voxels within the renal cortex. These differences also affect parameter ranges. Additionally, influences of noise and motion artifacts are assumed to have a more prominent role in in vivo DWI-parameter estimations. We tried to account for this in the simulations by using Rician noise with SDs of $0.01 \cdot S_0$, resulting in SNRs that were lower compared to in vivo measurements.

The main advantage of acquiring DWI-images with the accelerated sequence is a nine-fold reduction in acquisition time. While the parameter ranges are wider for f and D_p , these were similar for all other DWI-parameters when comparing the results of the accelerated sequence with the standard DWI-sequence. The images obtained with the accelerated sequence have higher SNRs compared to those of the standard sequence. This can be caused by image acquisition settings, such as the parallel imaging factor, voxel size, number of repetitions and bandwidth per pixel. Furthermore, differences in SNR can be a result of application of correction algorithms which were available in Siemens’ WIP protocol. In-plane image registration, filtering and denoising compensate motion effects. Application of these correction algorithms resulted in parameter maps that provide a better representation of the renal anatomy, in particularly in the D_p -map. However, this did not result in smaller parameter ranges. Therefore, future research should show whether the wider ranges in f and D_p outweigh the decrease in acquisition time, taking into account the clinical usability of DWI-sequences.

The main limitation of the validation study is the number of included volunteers. A broader study population could provide a more representative parameter estimation range. This has to be evaluated in further studies. Additionally, we only optimized the b -value distribution by means of computer simulations. A better optimization would require comparing numerous b -value sets using in vivo data. However, this would require impractical long acquisition times. Therefore, we only evaluated the optimized b -value distribution in DWI-parameter estimation and compared our results with the results of other studies. Additionally, we studied the influence of SMS imaging on the accuracy and precision of DWI-parameter estimation and acquisition times. However, the influence of other sequence parameters, like the echo time, should be studied for optimization of renal DWI, to provide even higher accuracy and precision in DWI-parameter estimation. Lastly, we only studied two fitting algorithms. These algorithms were evaluated since they are applied by most studies concerning IVIM model fitting. Other algorithms might improve robustness of fitting. Though, the long computational times of more advanced algorithms are detrimental for clinical use.

3.5 Conclusion

In conclusion, this study provides optimal b -value distributions for the IVIM and stretched exponential DWI-models acquired by computer simulations. This optimized distribution, containing b -values of 0, 30, 60, 110, 140, 800 and

1400 s/mm², was tested in healthy volunteers. In simulations, as well as in vivo, the Bayesian probability algorithm provided more precise parameter estimations for the IVIM model, compared to the TRR algorithm. There were no evident differences between the algorithms for the stretched exponential model. Lastly, the DWI-sequence using SMS imaging provided a nine-fold reduction in acquisition time, while the variability in parameters reflecting tissue diffusion were similar to those obtained with the standard DWI-sequence.

CHAPTER 4 – Diagnostic utility of multi-exponential diffusion-weighted imaging parameters in differentiating solid renal tumors

4.1 Introduction

In the previous chapter we evaluated the use of multi-exponential DWI-models in describing diffusion within the renal cortex, using an optimized b -value distribution. In this chapter, the usefulness of IVIM and stretched exponential DWI-parameters in differentiating renal tumors is studied.

As described earlier, nowadays, more than 60% of the RCC are detected incidentally in patients not suspected for renal cancer. This increase in incidence presents a challenge in diagnosis and management of renal tumors. In many cases, CT and MRI protocols that are used in daily clinical practice cannot accurately differentiate benign and malignant renal tumors. Consequently, in up to 20% of the cases benign histology is found after tumor nephrectomy. Additionally, RCC comprises several subtypes among which ccRCC, pRCC and chRCC are most prevalent. Disease prognosis varies widely between these histological subtypes. Therefore, with an aging patient population, discrimination between indolent and aggressive malignant renal tumors becomes more important, considering that treating early stage renal tumors with (partial) nephrectomy does not always improve patient health outcomes.[73] Furthermore, differentiation between ccRCC and other RCC subtypes is important since the subtypes show different sensitivities to systemic therapies.[74] Accordingly, non-invasive evaluation of solid renal tumors is of growing value in appropriate treatment selection.

Tissue diffusion and perfusion are influenced by tumor growth and may therefore be useful for lesion characterization. DWI can provide quantitative information about these changes. Because kidneys are highly vascularized organs, signal decay in DWI is expected to be influenced by factors other than pure diffusion. The mono-exponential model cannot properly describe these factors, which results in large ranges in ADC values within and overlap between tumor types. Therefore, reliable RCC subtyping and differentiation between benign and malignant renal tumors are hampered. Multi-exponential models are thought to provide an improved description of the signal decay and therefore provide more adequate diffusion parameters.

At histopathology, the RCC subtypes show differences in cellularity and vascularity. For example, ccRCC tend to have a more heterogenous texture compared to pRCC.[75] Therefore analysis of heterogeneity in multi-exponential DWI parameters within the lesions might be useful for tumor subtyping. Furthermore, tumor development is accompanied by changes in cell and tissue organization. These changes may differ between the tumor types. Therefore, the tumor induced changes in histogram parameters will be studied as well.

We hypothesize that multi-exponential DWI-parameters provide narrower ranges for tissue diffusion components compared to the mono-exponential model, which will improve differentiation of benign and malignant renal masses and RCC subtyping.

The purpose of this study was to determine if histogram analysis of IVIM and stretched exponential DWI-parameters can be used for histologic subtyping of RCC and differentiation of benign and malignant renal masses.

4.2 Methods

4.2.1 Patients

This prospective study was institutional review board approved and Health Insurance Portability and Accountability Act compliant. Written informed consent was obtained from all patients before enrollment. Seven consecutive patients were included between April 2019 and June 2019. Inclusion criteria were: patients diagnosed with a solid renal mass < 10 cm, suspected for RCC based on previous CT, MRI or ultrasonography examinations; and at least 18 years of age. Patients were excluded using the following criteria: contraindications for MRI, prior treatment for renal malignancies, artifacts in DW-images impeding reliable parameter calculation, and unconfirmed renal tumor type at histopathology.

4.2.2 MR Imaging

Pre-operative MR imaging was performed with a 3T MRI scanner (Magnetom prisma, Siemens Healthcare, Erlangen, Germany) and an 18-channel body phased-array coil. The scanning protocol consisted of the following sequences: an axial T2-weighted fat-suppressed HASTE in breath-hold and a coronal T2-weighted fat-suppressed turbo spin echo in expiration. For DWI, a respiratory triggered SS-EPI sequence was used in the coronal plane. The

following eight b -values were used: 0, 30, 60, 110, 140, 400, 800, and 1040 s/mm². Images with $b=0$ to 140 s/mm² were acquired with two averages and higher b -values were acquired with four averages. Diffusion gradients were applied in three directions. Respiratory triggering was performed using a liver dome scout. Additional imaging parameters can be found in Table 8. During reconstruction of the trace DW-images, motion correction consisting of in-plane registration, filtering, denoising and rescaling was performed (Siemens WIP prototype sequence no. 996C). The mean DWI acquisition time was 2:54 min (range: 2:15 to 3:50 min.).

4.2.2 Image analysis

Acquired images were transferred to an offline work-station for processing. Per patient, two 3-dimensional regions of interest (ROI) were drawn by one observer (with 1 year of experience in abdominal imaging), blinded for histopathological information using FireVoxel. The first ROI covered the entire tumor, while the second ROI covered a volume > 10 cm³ of renal cortex of the contralateral kidney. The tumor volumes were calculated using these ROIs.

To reduce partial volume effects, edge voxels of the tumor ROI were removed using an erosion method. IVIM and stretched exponential models were fitted to the signal intensities within both ROIs using MATLAB. For fitting, the BP algorithm was used. Following voxel-wise fitting, histograms of the diffusion parameters were generated. Outliers of the diffusion parameters were detected and removed using the interquartile range method. The following histogram parameters were analyzed: mean, mode, standard deviation, entropy, skewness, kurtosis and 5th, 10th, 25th, 75th, and 95th percentiles. Furthermore, tumor induced differences in histogram metrics between tumor and contralateral cortex tissue were calculated for all patients by calculating the difference in histogram parameters.

4.2.3 Histologic analysis

The partial nephrectomy specimens were fixed in formalin, processed and cut into 4 µm sections and stained with H&E. The slides were analyzed by board-certified pathologists, reporting the histologic tumor (sub)type, grade according to the WHO grading system, and presence or absence of hemorrhage, and necrosis.

4.2.4 Statistical analysis

The Mann-Whitney test was used to compare multi-exponential DWI-parameters between benign and malignant tumors. For pairwise comparisons between all tumor subtypes, independent-sample unequal variance 2-tailed t tests were performed. Tukey's range test was used to correct for multiple comparisons. A logistic regression model was created using parameters that were statistically significantly different between the groups. Collinearity of model parameters was tested using the VIF . A $VIF > 4$ was considered as presence of multicollinearity, and such parameters were excluded from the model.

ROC curve analyses were performed to test the ability of the diffusion parameters to differentiate the tumor subtypes and the AUCs were calculated. The optimal threshold was determined by evaluating the sensitivity, specificity and Youden index at different cutoff points. All statistical analyses were performed using SPSS (version 25.0, IBM-SPSS, Armonk, New York, USA). For all tests, a p -value ≤ 0.05 was considered statistically significant.

4.3 Results

Since inclusion is still ongoing, preliminary results are provided.

4.3.1 Patient characteristics

Up until now, MRI was performed and histological confirmation was obtained in six patients. Data of 1 patient was excluded due to presence of severe susceptibility artifacts within the tumor. The cohort included five patients (3 females, 2 males, mean age 64 years, range: 56-77) with 3 ccRCCs and 2 pRCCs. Mean tumor volumes were 37 and 120 cm³ for ccRCC and pRCC, respectively. (Table 9)

4.3.2 Subtype comparison

In Figure 9, examples of parameter maps of multi-exponential DWI-parameters are shown. The histogram parameters of all tumors are listed in Table 10. Only the 75th and 95th percentile values of D_t show no overlap between ccRCC and pRCC in both central tendency measure. When considering the mode as central tendency measure, the mode of DDC does not overlap between the RCC subtypes. The percentiles of D_t as well as the mode

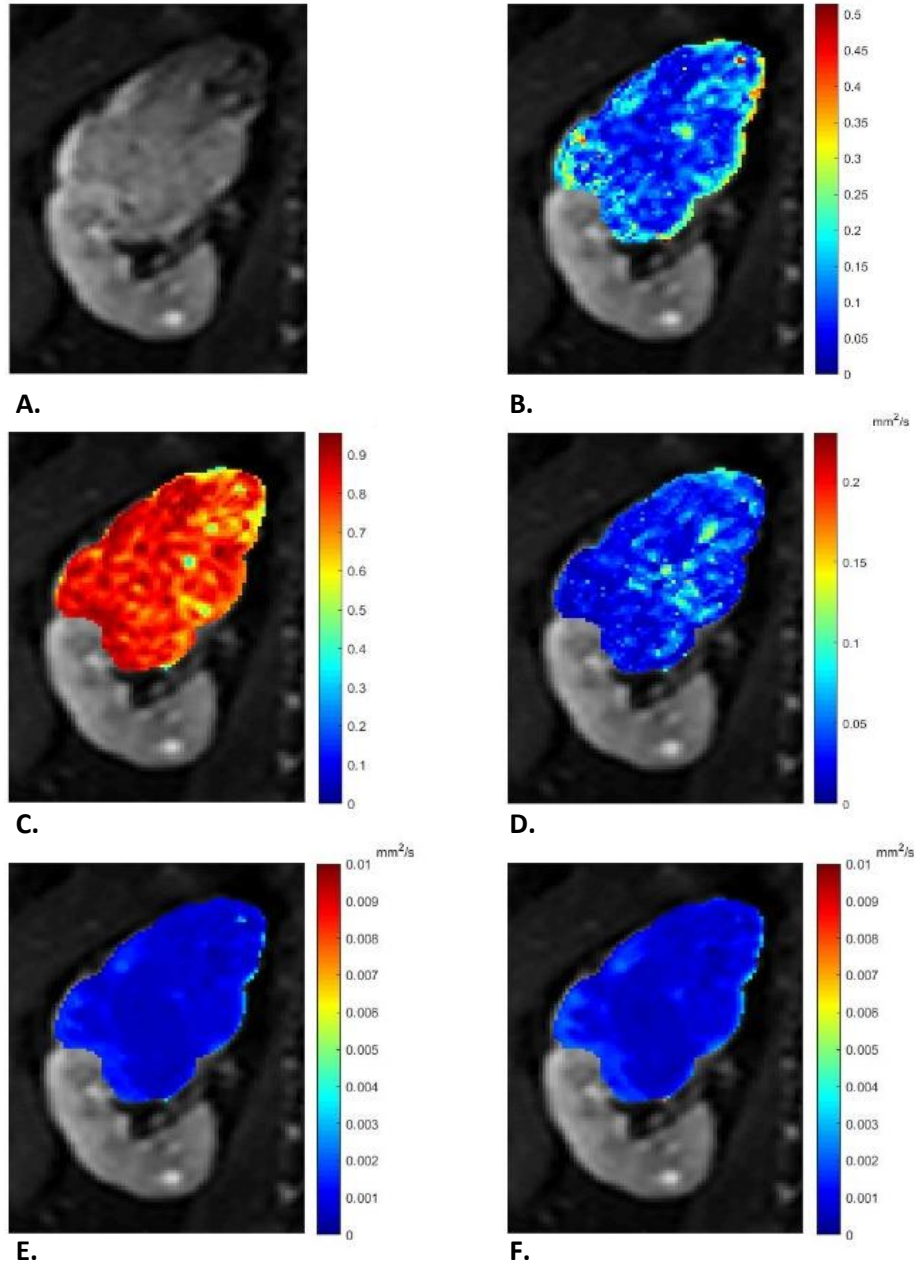


Figure 9. Parameter maps of multi-exponential DWI models of a papillary renal cell carcinoma. In A the b_0 -image is shown. Parameter maps of IVIM parameters f , D_p and D_t are shown in B, D and F, respectively. Parameter maps of stretched exponential parameters α and DDC are shown in C and E, respectively.

Table 8. Magnetic resonance imaging parameters

Parameter	Axial T2-weighted HASTE	Coronal T2-weighted Turbo Spin Echo	DWI
No. of slices	30	35	35
TR (ms)	2000	2750	2100
TE (ms)	93	104	51
Flip angle (°)	150	140	90
FOV (mm x mm)	450 x 450	400 x 400	308 x 380
Voxel size (mm)	1.4 x 1.4 x 4	1.3 x 1.3 x 5	1.5 x 1.5 x 5
Bandwidth (Hz/Px)	781	260	2440

of DDC is higher in ccRCC compared to pRCC. Other DWI-parameters demonstrate overlapping ranges. Boxplots of all parameters are shown in Appendix C.

In Table 11 the tumor induced changes in DWI parameter are shown. Histogram parameters that showed no overlap between the RCC subtypes were found for f , D_t , DDC and α . The mean, mode and 75th and 95th percentiles of f show a larger decrease for pRCC compared to ccRCC. Furthermore, the skewness in f remained the same for ccRCC, while pRCC show a higher skewness, compared to contralateral cortical tissue. The kurtosis in f was higher in pRCC and lower in ccRCC compared to the cortical tissue. The kurtosis in D_t and DDC decreased in pRCC, while it remained similar for ccRCC. The mode of the DDC histogram decreased more in pRCC compared to ccRCC. Finally, the mean of α decreased more for ccRCC than for pRCC. Boxplots of all parameter are shown in Appendix D.

4.3.3 Statistical analysis

A larger study population is required for statistical analysis.

4.4 Discussion

Reliable characterization of renal tumors is vital for appropriate treatment planning. Differentiation of benign and malignant tumors essential: benign tumors seldomly interfere with the patients' life expectancy and therefore invasive treatment involves unnecessary surgical risks. Based on currently available data, this study cannot perform this analysis. However, subtyping of renal malignancies is becoming increasingly important in evaluating treatment options. The subtypes differ in aggressiveness and likelihood of metastasis. In an aging population, with an increasing number of comorbidities, alternative treatments are preferred over nephrectomies in less aggressive tumors. Furthermore, in advanced disease, the subtypes show different sensitivities to targeted immunotherapies.

The preliminary results of this studies show overlapping ranges for ccRCC and pRCC in most of the DWI-parameters. Histogram parameters that do not overlap between the RCC subtypes are the 75th and 95th percentiles of D_t , and the mode of DDC . Prior studies showed that the mean of D_t was significantly different between ccRCC and pRCC.[76]. However, the values of D_t did overlap between both subtypes in their study as well.

There were differences in DWI-parameters obtained by using the mean and mode as central tendency measures for the estimates provided by the BP-algorithm. These differences appear mainly in histogram analyses of IVIM parameters f and D_p , but also occur in histogram parameter of D_t and α . Despite these differences, there seems to be no real difference in tumor differentiation ability between both measures.

Table 9. Patient demographics and tumor characteristics

	ccRCC (n=3)	pRCC (n=2)
Patient age (y)	64 (61 – 65)	67 (56 – 77)
Female	2 (66.7)	1 (50)
Body mass index (kg/m ²)	23.7 (23.3 – 24.3)	32.7 (29.6 – 35.9)
Tumor volume (cm ³)	37 (3 – 88)	120 (45 – 196)
Affected side		
Left	2 (66.6)	0
Right	1 (33.3)	2 (100)
Type of procedure		
Partial nephrectomy	1 (66.6)	2 (100)
Radical nephrectomy	2 (33.3)	0
T stage		
T1	2 (66.6)	1 (50)
T2	0	1 (50)
T3	1 (33.3)	0
T4	0	0
Nuclear grading (Fuhrman)		
I	0	0
II	1 (33.3)	2 (100)
III	2 (66.6)	0
IV	0	0

Data are presented as mean (range) or No. (%); NA = not applicable.

Table 10. Results of histogram analysis obtained with the Bayesian probability algorithm with mean as central tendency measure

		Mean	SD	Mode	Skewness	Kurtosis	Entropy
f	ccRCC	19.5 ± 2.2	8.1 ± 2.5	17.1 ± 2.1	0.72 ± 0.20	0.29 ± 0.37	3.34 ± 0.05
	pRCC	16.8 ± 7.7	8.1 ± 2.8	12.3 ± 5.1	0.87 ± 0.08	0.50 ± 0.19	3.30 ± 0.01
D_p	ccRCC	36.5 ± 12.7	23.5 ± 15.6	20.0 ± 9.2	1.03 ± 0.16	0.71 ± 0.33	3.32 ± 0.03
	pRCC	32.2 ± 0.03	17.0 ± 3.4	23.5 ± 8.3	0.93 ± 0.11	0.68 ± 0.05	3.32 ± 0.05
D_t	ccRCC	1.38 ± 0.11	0.37 ± 0.13	1.35 ± 0.18	0.42 ± 0.19	0.60 ± 0.33	2.58 ± 0.55
	pRCC	1.16 ± 0.24	0.37 ± 0.02	1.14 ± 0.43	0.60 ± 0.13	0.75 ± 0.98	2.93 ± 0.22
DDC	ccRCC	1.71 ± 0.15	0.48 ± 0.22	1.72 ± 0.04	0.42 ± 0.22	0.25 ± 0.23	3.21 ± 0.19
	pRCC	1.42 ± 0.46	0.47 ± 0.01	1.23 ± 0.73	0.41 ± 0.26	-0.05 ± 0.40	3.09 ± 0.16
α	ccRCC	0.79 ± 0.03	0.10 ± 0.04	0.87 ± 0.02	-0.96 ± 0.18	0.59 ± 0.14	3.30 ± 0.10
	pRCC	0.79 ± 0.05	0.10 ± 0.01	0.86 ± 0.07	-0.83 ± 0.28	0.20 ± 0.40	3.24 ± 0.01
		5 th percentile	10 th percentile	25 th percentile	75 th percentile	95 th percentile	
f	ccRCC	8.5 ± 4.3	10.1 ± 4.6	13.4 ± 3.9	24.2 ± 2.0	35.2 ± 5.1	
	pRCC	6.2 ± 3.7	7.8 ± 4.5	10.9 ± 5.7	21.5 ± 9.4	32.8 ± 13.0	
D_p	ccRCC	9.4 ± 3.1	12.4 ± 4.2	19.1 ± 4.8	48.4 ± 22.2	86.6 ± 46.7	
	pRCC	10.3 ± 2.3	13.4 ± 2.9	19.9 ± 3.5	40.9 ± 2.4	66.6 ± 7.6	
D_t	ccRCC	0.82 ± 0.25	0.92 ± 0.28	1.10 ± 0.25	1.61 ± 0.04	2.05 ± 0.13	
	pRCC	0.62 ± 0.15	0.73 ± 0.21	0.88 ± 0.30	1.43 ± 0.17	1.76 ± 0.17	
DDC	ccRCC	1.01 ± 0.42	1.12 ± 0.44	1.34 ± 0.39	2.00 ± 0.07	2.54 ± 0.15	
	pRCC	0.73 ± 0.27	0.87 ± 0.40	1.06 ± 0.52	1.79 ± 0.40	2.23 ± 0.50	
α	ccRCC	0.59 ± 0.11	0.64 ± 0.09	0.73 ± 0.05	0.87 ± 0.01	0.91 ± 0.02	
	pRCC	0.58 ± 0.06	0.63 ± 0.06	0.72 ± 0.06	0.87 ± 0.04	0.92 ± 0.03	

Data are presented as mean ± SD. Values for f are in percentage. D_p , D_t , and DDC are in $10^{-3} \text{ mm}^2/\text{s}$

Table 11. Results of histogram analysis of tumor induced differences, obtained with the Bayesian probability algorithm with mean as central tendency measure

		Mean	SD	Mode	Skewness	Kurtosis	Entropy
<i>f</i>	ccRCC	-5.3 ± 4.9	-0.2 ± 5.4	-6.8 ± 6.5	0.09 ± 0.25	-0.07 ± 0.19	0.09 ± 0.08
	pRCC	-6.0 ± 4.9	-1.7 ± 5.7	-5.1 ± 0.9	0.07 ± 0.13	-0.22 ± 0.14	0.00 ± 0.01
<i>D_p</i>	ccRCC	14.7 ± 11.7	12.0 ± 17.9	2.4 ± 8.9	-0.07 ± 0.29	-0.32 ± 0.67	0.01 ± 0.04
	pRCC	7.5 ± 4.7	2.5 ± 10.2	6.8 ± 8.1	0.03 ± 0.48	0.08 ± 0.30	-0.01 ± 0.05
<i>D_t</i>	ccRCC	-0.42 ± 0.32	0.13 ± 0.22	-0.44 ± 0.31	0.10 ± 0.90	-4.06 ± 4.63	0.49 ± 0.69
	pRCC	-0.51 ± 0.24	0.14 ± 0.05	-0.50 ± 0.41	1.33 ± 0.41	-1.62 ± 1.04	-0.11 ± 0.02
<i>DDC</i>	ccRCC	-0.60 ± 0.14	0.23 ± 0.31	-0.49 ± 0.26	-0.26 ± 0.28	-0.19 ± 0.61	0.03 ± 0.25
	pRCC	-0.73 ± 0.54	0.25 ± 0.07	-0.87 ± 0.70	0.12 ± 0.36	-0.39 ± 0.56	0.25 ± 0.10
<i>α</i>	ccRCC	-0.04 ± 0.08	0.04 ± 0.06	0.02 ± 0.06	-0.32 ± 0.38	0.45 ± 0.80	-0.01 ± 0.23
	pRCC	-0.05 ± 0.08	0.05 ± 0.02	0.01 ± 0.11	-0.12 ± 0.25	-0.26 ± 0.23	-0.05 ± 0.05
		5 th percentile	10 th percentile	25 th percentile	75 th percentile	95 th percentile	
<i>f</i>	ccRCC	-4.7 ± 3.0	-5.1 ± 2.5	-5.3 ± 2.5	-5.3 ± 4.6	-5.2 ± 15.3	
	pRCC	-3.6 ± 2.3	-4.0 ± 2.7	-4.6 ± 3.7	-7.1 ± 12.4	-8.3 ± 19.4	
<i>D_p</i>	ccRCC	1.3 ± 5.3	2.9 ± 6.2	6.0 ± 4.5	20.9 ± 23.1	40.3 ± 50.6	
	pRCC	3.6 ± 3.3	4.3 ± 4.1	5.7 ± 4.2	8.3 ± 10.0	11.3 ± 28.9	
<i>D_t</i>	ccRCC	-0.60 ± 0.47	-0.58 ± 0.45	-0.55 ± 0.41	-0.32 ± 0.33	-0.13 ± 0.51	
	pRCC	-0.58 ± 0.05	-0.63 ± 0.11	-0.68 ± 0.28	-0.39 ± 0.22	-0.21 ± 0.28	
<i>DDC</i>	ccRCC	-0.96 ± 0.30	-0.90 ± 0.30	-0.78 ± 0.22	-0.47 ± 0.35	-0.23 ± 0.56	
	pRCC	-1.08 ± 0.31	-1.00 ± 0.42	-0.94 ± 0.54	-0.53 ± 0.51	-0.33 ± 0.72	
<i>α</i>	ccRCC	-0.14 ± 0.19	-0.11 ± 0.13	-0.06 ± 0.12	-0.01 ± 0.04	0.00 ± 0.03	
	pRCC	-0.15 ± 0.10	-0.13 ± 0.10	-0.08 ± 0.09	-0.01 ± 0.06	0.01 ± 0.04	

Data are presented as mean ± SD. Values for *f* are in percentage, *D_p*, *D_t*, and *DDC* are in ·10⁻³ mm²/s

The mean of the parameters that reflect tissue diffusion (D_t and DDC) were both lower in RCC compared to contralateral cortical renal tissue. This implies that both tumor types cause diffusion restriction. However, the tumor induced difference is similar for both RCC subtypes. A decrease in kurtosis of D_t and DDC seen as well, which implies that more distinct diffusion rates occur in tumor tissue compared to normal cortical renal tissue. The kurtosis in D_t and DDC , as well as the mode of DDC , decrease more in pRCC than in ccRCC. The decrease in mode reflects a drop in number of high DDC values, since the mean does not change considerably.

A smaller pseudodiffusion fraction, f , is also found in tumorous tissue. The decrease is larger for pRCC than for ccRCC, which can be seen in the mean, mode and 75th and 95th percentiles of the histogram values of f . Thus, the effect of pseudodiffusion to signal decay within a voxel decreases. The kurtosis in pRCC rise as well, which implies that the pseudodiffusion fractions become centered around one value. However, the skewness of f increases in pRCC, which means that within some of the voxels large pseudodiffusion fractions are still present.

The preliminary results of this studies show overlapping ranges for ccRCC and pRCC in most of the DWI-parameters. Histogram parameters that not overlap between the RCC subtypes are the 75th and 95th percentiles of D_t , and the mode of DDC . Prior studies showed that the mean of D_t was significantly different between ccRCC and pRCC.[76]. However, the values of D_t did overlap between both subtypes in their study as well.

There were differences in DWI-parameters obtained by using the mean and mode as central tendency measures for the estimates provided by the BP-algorithm. However, there seems to be no real difference in tumor differentiation ability between both measures.

The mean of the parameters that reflect diffusion (D_t and DDC) were both lower in RCC compared to contralateral cortical renal tissue. This implies that both tumor types cause diffusion restriction. However, the tumor induced difference is similar for both RCC subtypes. A decrease in kurtosis of D_t and DDC seen as well, which implies that more distinct diffusion rates occur in tumor tissue compared to normal cortical renal tissue. The kurtosis in D_t and DDC , as well as the mode of DDC , decrease more in pRCC than in ccRCC. The decrease in mode reflects a drop in number of high DDC values.

A smaller perfusion fraction, f , is also found in tumorous tissue. A larger decrease is seen for pRCC than for ccRCC, which can be seen in the mean, mode and 75th and 95th percentiles of the histogram values of f . Thus, the effect of pseudodiffusion to signal decay within a voxel diminishes. The kurtosis in pRCC rise as well, which implies that the perfusion fractions become centered around one value. However, the skewness of f increases in pRCC, which means that within some of the voxels, large perfusion fraction are still present.

The measure for intravoxel inhomogeneity, α , could also aid differentiation between pRCC and ccRCC. Both RCC subtypes have lower α , and therefore a higher degree of intravoxel inhomogeneity, compared to contralateral cortical renal tissue. This can be explained by an increase in histological heterogeneity, cellular pleomorphism, presence of micro-vessels, and presence of microscopic necrotic and cystic foci. ccRCC show a larger decrease compared to pRCC, which represents a stronger increase in intravoxel inhomogeneity in this tumor type.

Absolute differences in histogram parameters are small, thus a larger patient population is required to see if the results remain valid. Furthermore, more subtypes and benign renal tumors have to be included to evaluate the clinical value of the multi-exponential DWI-parameters.

4.5 Conclusion

Most of the multi-exponential DWI-parameters show overlapping values between pRCC and ccRCC. A larger patient population should show if there are any statistically significant differences between RCC subtypes. Furthermore, the analysis of tumor induced changes shows promising results for differentiation of the subtypes.

CHAPTER 5 – General conclusion, discussion and future perspectives

The goal of this thesis was to evaluate the role of diffusion-weighted imaging parameters in differentiation of solid renal tumors. The first sub-question that was formulated was: *To what extent can the combination of mono-exponential diffusion-weighted imaging parameters and patient characteristics differentiate oncocytoma from renal cell carcinoma?* It can be concluded that DWI analysis using mono-exponential model parameters shows promising results in differentiation of oncocytoma and renal cell carcinoma. While mean *ADC* values did not aid this differentiation, *ADC* histogram parameters were informative. However, the wide range in parameters hampers reliable differentiation and therefore the clinical value may be limited.

Subsequently, the following sub-question was studied: *What b-value distribution and fitting algorithms provide a robust calculation of multi-exponential diffusion-weighted imaging parameters?* By performing computer simulations, optimized b-value distributions were obtained for the intravoxel incoherent motion and stretched exponential models. Furthermore, for the IVIM model, the Bayesian probability algorithm provided a more robust parameter estimation compared to the trust-region reflective algorithm. For the stretched exponential model, there was no pronounced difference between the algorithms. These results were confirmed by validation measurements in healthy volunteers. Using an accelerated DWI sequence instead of a standard DWI-sequence resulted in images with a higher signal-to-noise ratio and a nine-fold reduction in acquisition time. The higher image quality obtained in a much shorter scan time is beneficial for using DWI with multiple b-values in clinical practice.

The results of the third chapter were used in a study to answer the following sub-question: *What is the ability of multi-exponential diffusion-weighted imaging parameters in subtyping of renal tumors?* IVIM and stretched exponential histogram parameters show promising results for differentiation of clear cell and papillary renal cell carcinoma. Furthermore, adding information about tumor-induced differences seems to be of added value for renal tumor differentiation. Research should be continued to show if these results are valid in a larger study population and to see if differentiation of other subtypes, as well as differentiation of benign and malignant renal tumors, can be improved by these parameters. In the current studies, first-order texture metrics, derived from DWI-parameter histograms were used to describe tumor heterogeneity. However, using this method, spatial information about the parameter values is lost. The value of higher order texture metrics that could provide this information should be studied.

To conclude, the main research question was: *What is the role of diffusion-weighted imaging in the differentiation of solid renal masses?* The research conducted in this thesis shows that DWI has a promising role in this differentiation. Although further research is needed, the use of DWI can improve the differentiation of renal tumor types and therefore improve clinical decision making concerning treatment options in patients with renal tumors. Consequently, it may help reduce the number of superfluous (partial) nephrectomies.

References

- [1] B. Ljungberg *et al.*, “EAU guidelines on renal cell carcinoma,” *Eur. Assoc. Urol.*, 2018.
- [2] H. Moch, A. L. Cubilla, P. A. Humphrey, V. E. Reuter, and T. M. Ulbright, “The 2016 WHO classification of tumours of the urinary system and male genital organs—Part A: Renal, penile, and testicular tumours,” *Eur. Urol.*, vol. 70, no. 1, pp. 93–105, 2016.
- [3] J. C. Cheville, C. M. Lohse, H. Zincke, A. L. Weaver, and M. L. Blute, “Comparisons of outcome and prognostic features among histologic subtypes of renal cell carcinoma,” *Am. J. Surg. Pathol.*, vol. 27, no. 5, pp. 612–624, 2003.
- [4] A. Kutikov *et al.*, “Incidence of benign pathologic findings at partial nephrectomy for solitary renal mass presumed to be renal cell carcinoma on preoperative imaging,” *Urology*, vol. 68, no. 4, pp. 737–740, 2006.
- [5] S. Choudhary, A. Rajesh, N. J. Mayer, K. A. Mulcahy, and A. Haroon, “Renal oncocytoma: CT features cannot reliably distinguish oncocytoma from other renal neoplasms,” *Clin. Radiol.*, vol. 64, no. 5, pp. 517–522, 2009.
- [6] K. Sasaguri and N. Takahashi, “CT and MR imaging for solid renal mass characterization,” *European Journal of Radiology*. 2018.
- [7] H. Tanaka *et al.*, “Diffusion-weighted magnetic resonance imaging in the differentiation of angiomyolipoma with minimal fat from clear cell renal cell carcinoma,” *Int. J. Urol.*, vol. 18, no. 10, pp. 727–730, 2011.
- [8] T. J. van Oostenbrugge, J. J. Fütterer, and P. F. A. Mulders, “Diagnostic Imaging for Solid Renal Tumors: A Pictorial Review,” *Kidney Cancer*, vol. 2, no. 2, pp. 79–93, 2018.
- [9] B. A. Shannon, R. J. Cohen, H. de Bruto, and R. J. Davies, “The value of preoperative needle core biopsy for diagnosing benign lesions among small, incidentally detected renal masses,” *J. Urol.*, vol. 180, no. 4, pp. 1257–1261, 2008.
- [10] H. Li *et al.*, “Whole-tumor quantitative apparent diffusion coefficient histogram and texture analysis to differentiation of minimal fat angiomyolipoma from clear cell renal cell carcinoma,” *Acad. Radiol.*, vol. 26, no. 5, pp. 632–639, 2018.
- [11] Y. Mytsyk *et al.*, “Renal cell carcinoma: applicability of the apparent coefficient of the diffusion-weighted estimated by MRI for improving their differential diagnosis, histologic subtyping, and differentiation grade,” *Int. Urol. Nephrol.*, vol. 49, no. 2, pp. 215–224, 2017.
- [12] V. Baliyan, C. J. Das, R. Sharma, and A. K. Gupta, “Diffusion weighted imaging: Technique and applications,” *World J. Radiol.*, vol. 8, no. 9, pp. 785–798, 2016.
- [13] N. S. White *et al.*, “Diffusion-weighted imaging in cancer: Physical foundations and applications of restriction spectrum imaging,” *Cancer Res.*, vol. 74, no. 17, pp. 4638–4652, 2014.
- [14] A. R. Padhani *et al.*, “Diffusion-weighted magnetic resonance imaging as a cancer biomarker: consensus and recommendations,” *Neoplasia*, vol. 11, no. 2, pp. 102–125, 2009.
- [15] J. L. Zhang *et al.*, “Variability of renal apparent diffusion coefficients: Limitations of the monoexponential model for diffusion quantification,” *Radiology*, vol. 254, no. 3, pp. 783–792, 2010.
- [16] H. Wang and X. Zhang, “Renal Cell Carcinoma : Diffusion- weighted MR Imaging for Subtype Purpose : Methods : Results : Conclusion :,” *Radiology*, vol. 257, no. 1, pp. 135–143, 2010.
- [17] B. Paudyal *et al.*, “The role of the ADC value in the characterisation of renal carcinoma by diffusion-weighted MRI,” *Br. J. Radiol.*, vol. 83, no. 988, pp. 336–343, 2010.
- [18] Y. Li, Y. Wang, J. Qin, J. Wu, X. Dai, and J. Xu, “Meta-analysis of diffusion-weighted imaging in the

- differential diagnosis of renal lesions," *Clin. Imaging*, vol. 52, no. 100, pp. 264–272, 2018.
- [19] A. Li *et al.*, "Subtype differentiation of small (≤ 4 cm) solid renal mass using volumetric histogram analysis of DWI at 3-T MRI," *AJR. Am. J. Roentgenol.*, vol. 211, no. 3, pp. 614–623, 2018.
 - [20] B. Gaing *et al.*, "Subtype differentiation of renal tumors using voxel-based histogram analysis of intravoxel incoherent motion parameters," *Invest. Radiol.*, vol. 50, no. 3, 2015.
 - [21] R. L. Siegel, K. D. Miller, and A. Jemal, "Cancer statistics, 2018," *CA. Cancer J. Clin.*, vol. 67, no. 1, pp. 7–30, 2017.
 - [22] G. Novara *et al.*, "Updated systematic review and meta-analysis of the comparative data on colposuspensions, pubovaginal slings, and midurethral tapes in the surgical treatment of female stress urinary incontinence," *Eur. Urol.*, vol. 58, no. 2, pp. 218–238, 2010.
 - [23] A. B. Rosenkrantz, N. Hindman, E. F. Fitzgerald, B. E. Niver, J. Melamed, and J. S. Babb, "MRI Features of Renal Oncocytoma and Chromophobe Renal Cell Carcinoma," *AJR. Am. J. Roentgenol.*, 2010.
 - [24] J. H. Kim, H. Y. Sun, J. Hwang, S. S. Hong, Y. J. Cho, and S. W. Doo, "Diagnostic accuracy of contrast-enhanced computed tomography and contrast-enhanced magnetic resonance imaging of small renal masses in real practice : sensitivity and specificity according to subjective radiologic interpretation," *World J. Surg. Oncol.*, pp. 1–8, 2016.
 - [25] S. R. Prasad, V. R. Surabhi, C. O. Menias, A. A. Raut, and K. N. Chintapalli, "Benign renal neoplasms in adults: Cross-sectional imaging findings," *Am. J. Roentgenol.*, vol. 190, no. 1, pp. 158–164, 2008.
 - [26] S. G. Silverman, G. M. Israel, and Q.-D. Trinh, "Incompletely characterized incidental renal masses: emerging data support conservative management," *Radiology*, vol. 275, no. 1, pp. 28–42, 2015.
 - [27] G. Kovacs, M. Akhtar, Bruce J. Beckwith, P. Bugert, Colin S. Cooper, and B. Delahunt, "The heidelberg classification of renal cell tumours," *J. Pathol.*, vol. 183, pp. 131–133, 1997.
 - [28] S. Kawaguchi, K. A. Fernandes, A. Finelli, M. Robinette, N. Fleshner, and M. A. S. Jewett, "Most renal oncocytomas appear to grow: Observations of tumor kinetics with active surveillance," *J. Urol.*, vol. 186, no. 4, pp. 1218–1222, 2011.
 - [29] A. M. Hotker *et al.*, "Differentiation of clear cell renal cell carcinoma from other renal cortical tumors by use of a quantitative multiparametric MRI approach," *Am. J. Roentgenol.*, vol. 208, no. 3, pp. W85–W91, 2017.
 - [30] G. Giannarini, G. Petralia, and H. C. Thoeny, "Potential and limitations of diffusion-weighted magnetic resonance imaging in kidney, prostate, and bladder cancer including pelvic lymph node staging: A critical analysis of the literature," *Eur. Urol.*, vol. 61, no. 2, pp. 326–340, 2012.
 - [31] E. A. Lassel, R. Rao, C. Schwenke, S. O. Schoenberg, and H. J. Michaely, "Diffusion-weighted imaging of focal renal lesions: A meta-analysis," *Eur. Radiol.*, vol. 24, no. 1, pp. 241–249, 2014.
 - [32] H. Chandarana *et al.*, "Diffusion-weighted intravoxel incoherent motion imaging of renal tumors with histopathologic correlation," *Invest. Radiol.*, vol. 47, no. 12, pp. 688–696, 2012.
 - [33] Y. D. Zhang *et al.*, "Comparison of utility of histogram apparent diffusion coefficient and R2* for differentiation of low-grade from high-grade clear cell renal cell carcinoma," *Am. J. Roentgenol.*, vol. 205, no. 2, pp. W193–W201, 2015.
 - [34] P. M. Pierorazio *et al.*, "Management of renal masses and localized renal cancer: Systematic review and meta-analysis," *J Urol.*, vol. 196, no. 4, pp. 989–999, 2016.
 - [35] "Clinical practice guideline for the evaluation and management of chronic kidney disease," *KDIGO*, vol. 3, no. 1, 2013.
 - [36] D. C. Hoaglin and B. Iglewicz, "Fine-tuning some resistant rules for outlier labeling," *J. Am. Stat. Assoc.*, vol. 82, no. 400, pp. 1147–1149, 1987.

- [37] S. A. Waugh *et al.*, "Magnetic resonance imaging texture analysis classification of primary breast cancer," *Eur. Radiol.*, vol. 26, no. 2, pp. 322–330, 2016.
- [38] S. Wang *et al.*, "Texture analysis of diffusion weighted imaging for the evaluation of glioma heterogeneity based on different regions of interest," *Oncol. Lett.*, vol. 15, no. 5, pp. 7297–7304, 2018.
- [39] A. A. K. A. Razek, A. Farouk, A. Mousa, and M. Nancy Nabil, "Role of diffusion-weighted magnetic resonance imaging in characterization of renal tumors," *J. Comput. Assist. Tomogr.*, vol. 35, no. 3, pp. 332–336, 2011.
- [40] B. Taouli *et al.*, "Renal Lesions: Characterization with diffusion-weighted imaging versus contrast-enhanced MR imaging," *Radiology*, vol. 251, no. 2, 2009.
- [41] A. M. Doshi, W. C. Huang, N. M. Donin, and H. Chandarana, "MRI features of renal cell carcinoma that predict favorable clinicopathologic outcomes," *Am. J. Roentgenol.*, 2015.
- [42] B. Gaing *et al.*, "Subtype differentiation of renal tumors using voxel-based histogram analysis of intravoxel incoherent motion parameters," *Invest. Radiol.*, vol. 50, no. 3, pp. 144–152, 2015.
- [43] H. Chandarana *et al.*, "Histogram analysis of whole-lesion enhancement in differentiating clear cell from papillary subtype of renal cell cancer," *Radiology*, vol. 265, no. 3, pp. 790–798, 2012.
- [44] D. Le Bihan, E. Breton, D. Lallemant, and P. Grenier, "MR imaging of intravoxel incoherent motions: application to diffusion and perfusion in neurologic disorders," *Radiology*, vol. 161, no. 2, pp. 401–407, 1986.
- [45] S. Ichikawa, U. Motosugi, T. Ichikawa, K. Sano, H. Morisaka, and T. Araki, "Intravoxel incoherent motion imaging of the kidney: Alterations in diffusion and perfusion in patients with renal dysfunction," *Magn. Reson. Imaging*, vol. 31, no. 3, pp. 414–417, 2013.
- [46] B. Ebrahimi, N. Rihal, J. R. Woollard, J. D. Krier, A. Eirin, and L. O. Lerman, "Assessment of renal artery stenosis using intravoxel incoherent motion diffusion-weighted magnetic resonance imaging analysis," *Invest. Radiol.*, vol. 49, no. 10, pp. 640–646, 2014.
- [47] H. Chandarana, V. S. Lee, Elizabeth M. Hecht, B. Taouli, and E. E. Sigmund, "Comparison of biexponential and monoexponential model of diffusion weighted imaging in evaluation of renal lesions," *Invest. Radiol.*, vol. 46, no. 5, pp. 285–291, 2011.
- [48] S. van Baalen, A. Leemans, P. Dik, M. R. Lilien, B. ten Haken, and M. Froeling, "Intravoxel incoherent motion modeling in the kidneys: Comparison of mono-, bi-, and triexponential fit," *J. Magn. Reson. Imaging*, vol. 46, no. 1, pp. 228–239, 2017.
- [49] Y. Ding *et al.*, "Differentiating between malignant and benign renal tumors: do IVIM and diffusion kurtosis imaging perform better than DWI?," *Eur. Radiol.*, 2019.
- [50] K. M. Bennett, K. M. Schmainda, R. Bennett, D. B. Rowe, H. Lu, and J. S. Hyde, "Characterization of continuously distributed cortical water diffusion rates with a stretched-exponential model," *Magn. Reson. Med.*, vol. 50, no. 4, pp. 727–734, 2003.
- [51] J. F. Kallehauge *et al.*, "Apparent Diffusion Coefficient (ADC) as a quantitative parameter in diffusion weighted MR imaging in gynecologic cancer: Dependence on b-values used," *Acta Oncol. (Madr.)*, vol. 49, no. 7, pp. 1017–1022, 2010.
- [52] D. M. Koh, D. J. Collins, and M. R. Orton, "Intravoxel incoherent motion in body diffusion-weighted MRI: Reality and challenges," *Am. J. Roentgenol.*, vol. 196, no. 6, pp. 1351–1361, 2011.
- [53] H. Li *et al.*, "Monoexponential, biexponential, and stretched exponential diffusion-weighted imaging models: Quantitative biomarkers for differentiating renal clear cell carcinoma and minimal fat angiomyolipoma," *J. Magn. Reson. Imaging*, 2017.
- [54] J. Zhang *et al.*, "Comparison of monoexponential, biexponential, stretched-exponential, and kurtosis models of diffusion-weighted imaging in differentiation of renal solid masses," *Korean J. Radiol.*, vol. 20, no. 5, p. 791, 2019.

- [55] O. J. Gurney-Champion *et al.*, "Comparison of six fit algorithms for the intravoxel incoherent motion model of diffusionweighted magnetic resonance imaging data of pancreatic cancer patients," *PLoS One*, vol. 13, no. 4, pp. 1–18, 2018.
- [56] G. Y. Cho *et al.*, "Comparison of fitting methods and b-value sampling strategies for intravoxel incoherent motion in breast cancer," *Magn. Reson. Med.*, 2015.
- [57] O. Gustafsson, M. Montelius, G. Starck, and M. Ljungberg, "Impact of prior distributions and central tendency measures on Bayesian intravoxel incoherent motion model fitting," *Magn. Reson. Med.*, vol. 79, no. 3, pp. 1674–1683, 2018.
- [58] R. Van de Schoot, D. Kaplan, J. Denissen, J. B. Asendorpf, F. J. Neyer, and M. A. G. van Aken, "A gentle introduction to Bayesian analysis: applications to developmental research," *Child Dev.*, vol. 85, no. 3, pp. 842–860, 2014.
- [59] M. R. Orton, D. J. Collins, D.-M. Koh, and M. O. Leach, "Improved intravoxel incoherent motion analysis of diffusion weighted imaging by data driven Bayesian modeling," *Magn. Reson. Med.*, vol. 71, no. 1, pp. 411–20, 2014.
- [60] S. Barbieri, O. F. Donati, J. M. Froehlich, and H. C. Thoeny, "Impact of the calculation algorithm on biexponential fitting of diffusion-weighted MRI in upper abdominal organs," *Magn. Reson. Med.*, vol. 75, no. 5, pp. 2175–2184, 2016.
- [61] P. M. Robson, A. K. Grant, A. J. Madhuranthakam, R. Lattanzi, D. K. Sodickson, and C. A. McKenzie, "Comprehensive quantification of signal-to-noise ratio and g-factor for image-based and k-space-based parallel imaging reconstructions," *Magn Reson Imaging*, vol. 60, no. 4, pp. 895–907, 2008.
- [62] V. M. Runge, J. K. Richter, and J. T. Heverhagen, "Speed in Clinical Magnetic Resonance," *Invest. Radiol.*, vol. 52, no. 1, pp. 1–17, 2017.
- [63] F. A. Breuer, M. Blaimer, R. M. Heidemann, M. F. Mueller, M. A. Griswold, and P. M. Jakob, "Controlled aliasing in parallel imaging results in higher acceleration (CAIPIRINHA) for multi-slice imaging," *Magn. Reson. Med.*, vol. 53, no. 3, pp. 684–691, 2005.
- [64] J. Pan *et al.*, "Measurement and scan reproducibility of parameters of intravoxel incoherent motion in renal tumor and normal renal parenchyma: a preliminary research at 3.0 T MR," *Abdom. Radiol.*, vol. 43, no. 7, pp. 1739–1748, 2018.
- [65] H. Li *et al.*, "Monoexponential, biexponential, and stretched exponential diffusion-weighted imaging models: Quantitative biomarkers for differentiating renal clear cell carcinoma and minimal fat angiomyolipoma," *J. Magn. Reson. Imaging*, vol. 46, no. 240–247, 2017.
- [66] J. L. Zhang *et al.*, "Optimization of b-value sampling for diffusion-weighted imaging of the kidney," *Magn. Reson. Med.*, 2012.
- [67] P. T. While, "A comparative simulation study of bayesian fitting approaches to intravoxel incoherent motion modeling in diffusion-weighted MRI," *Magn. Reson. Med.*, vol. 78, no. 6, pp. 2373–2387, 2017.
- [68] O. Jalnefjord *et al.*, "Comparison of methods for estimation of the intravoxel incoherent motion (IVIM) diffusion coefficient (D) and perfusion fraction (f)," *Magn. Reson. Mater. Physics, Biol. Med.*, vol. 31, no. 6, pp. 715–723, 2018.
- [69] A. Lemke, B. Stieltjes, L. R. Schad, and F. B. Laun, "Toward an optimal distribution of b values for intravoxel incoherent motion imaging," *Magn. Reson. Imaging*, vol. 29, no. 6, pp. 766–776, 2011.
- [70] C. Ye *et al.*, "Estimation of intravoxel incoherent motion parameters using low b-values," *PLoS One*, vol. 14, no. 2, pp. 1–16, 2019.
- [71] J. Zhang *et al.*, "Quantification of renal diffusion-weighted images using a bi-exponential model," *Proc. Intl. Soc. Mag. Reson. Med*, 2009.
- [72] M. Notohamprodo *et al.*, "Combined intravoxel incoherent motion and diffusion tensor imaging of renal diffusion and flow anisotropy," *Magn. Reson. Med.*, 2015.

- [73] S. K. Kang, W. C. Huang, P. V. Pandharipande, and H. Chandarana, "Solid renal masses: What the numbers tell us," *Am. J. Roentgenol.*, vol. 202, no. 6, pp. 1196–1206, 2014.
- [74] A. J. Schrader, P. J. Olbert, A. Hegele, Z. Varga, and R. Hofmann, "Metastatic non-clear cell renal cell carcinoma: Current therapeutic options," *BJU Int.*, vol. 101, no. 11, pp. 1343–1345, 2008.
- [75] V. F. Muglia and A. Prando, "Renal cell carcinoma: histological classification and correlation with imaging findings," *Radiol. Bras.*, vol. 48, no. 3, pp. 166–174, 2015.
- [76] A. Li *et al.*, "Subtype differentiation of small (≤ 4 cm) solid renal mass using volumetric histogram analysis of DWI at 3-T MRI," *Am. J. Roentgenol.*, 2018.

Appendix A Design parameters for computer simulations

A.1 b -value distribution

In this study, b -values were sampled in steps of 30 s/mm² between b_0 and b_{800} , and in steps of 100 s/mm² between b_{900} and b_{1400} to obtain an initial b -value set of three b -values. When adding b -values to this set, b -values were sampled in steps of 10 s/mm² between b_0 and b_{800} , and in steps of 50 s/mm² between b_{850} and b_{1400} . Larger step sizes were used in higher b -values since pseudodiffusion effects, and therefore deviations from the mono-exponential model, appear mainly at low b -values. To describe these effects accurately, low b -values have to be sampled extensively. Ideally, all b -values should be sampled. However, this would require extensive computation time. Furthermore, the influence of smaller steps than 10 s/mm² on the optimized b -value distribution are expected to be small.

Furthermore, in finding the optimized b -value distribution, a set of 25 b -values was obtained. This number of b -values is not useful in clinical practice due to the long acquisition times. However, this number was used to study the decrease in fitting error per number of b -values. Thereby, an inflection point in decrease of error per added b -value could be found. This inflection point shows where the decrease in error outweighs the addition of adding another b -value to the distribution. Based on this point, an optimized b -value distribution was obtained.

A.2 Data noise

Prior studies have shown that the fitting errors depend on the SNR.[69] Higher SNRs provide better fitting results. In the current study, the influence of noise on fitting results was not studied. In simulation of signal intensities, Rician noise with an SD of $0.01 \cdot S_0$ was used. In this manner, the SNR in simulations was lower than the SNR in DWI-images obtained in healthy volunteers. Therefore, the influence of the SNR on fitting errors in DWI-images was expected to be smaller than in simulations. However, systematic errors, like partial volume effects and organ movement, were not taken into account in the simulations. These errors influence parameter fitting errors and could therefore also influence optimal parameter selection.

A.3 Fitting algorithms

In the current study, two fitting algorithms were evaluated in fitting multi-exponential models to signal intensities. The segmented fitting method was also considered since prior studies showed that this method facilitates robust data fitting. In this method, D_t is determined based on fitting the mono-exponential model to high b -values. This can be seen by the orange line in Figure A1. Subsequently, f is determined by taking the difference between the measured signal intensity at S_0 and the intersection of the mono-exponential fit and the y-axis (green line in Figure A1). Finally, D_p is determined by fitting the model to the data, using the previously determined values for D_t and

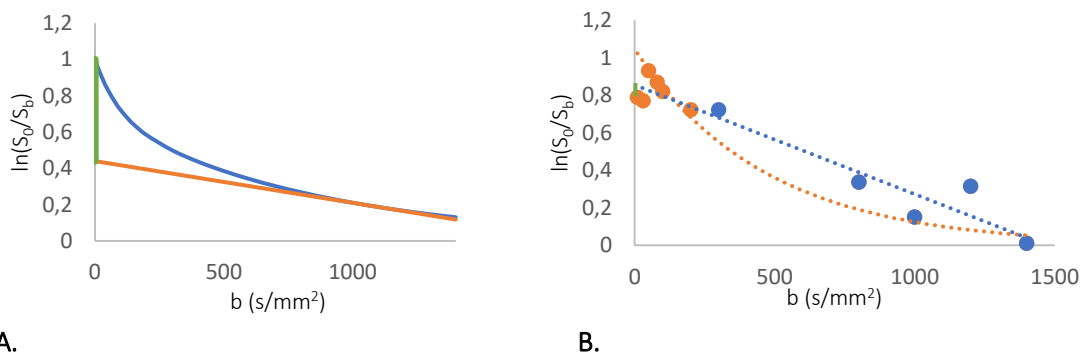


Figure A1. Segmented fitting method. D_t is determined based on the slope of the signal intensities at high b -values (orange). Subsequently, the intercept of this slope with the y-axis is used to determine f (green). Lastly, D_p is determined by fitting the intravoxel incoherent motion model to the data using the determined f and D_p . A problem arises when the intersect of the slope provides a higher value than the signal intensity measured at $b = 0$ s/mm², since this results in a negative f as can be seen in B. Now the results of the segmented fit are not clinically useful.

f . However, this method could provide erroneous measurements when the intersect of the mono-exponential fitting of high b -values is higher than the measured S_0 . This can happen due to influences of noise and systematic errors, as can be seen in Figure A1. Now f is negative, which is physiologically impossible. Since this study was focused on voxel-wise fitting and whole-tumor volumetric DWI-parameter estimation, the segmented model was not further studied.

A.4 Data fitting

In the BP algorithm, lognormal priors were used. The uniform prior is less informative than the other priors and therefore it has a small influence on the posterior probability of the parameter estimates. However, Gustafsson et al. (2018) studied the influence of uniform, lognormal and reciprocal priors in estimation of IVIM parameters.[57] They found that the uniform and lognormal priors provided small errors in estimations of the parameters, even at low SNRs. The bias in estimation of f and D_t was slightly smaller when the lognormal prior was used compared to the uniform prior, and the estimation of D_p was substantially better when using the lognormal prior. Therefore this prior was used in current study.

Table A1. Overview of diffusion-weighted imaging parameters used for simulations

	Ground truth parameters	Initial parameter estimates	Parameter constraints	Prior distributions in BP
IVIM				
f	30	10	$0 \leq f \leq 100$	Uniform
D_p	50	10	$0 \leq D_p \leq 1$	Lognormal ($\sigma=1, \mu = -6$)
D_t	1.2	1.0	$0 \leq D_t \leq 0.1$ $D_t < D_p$	Lognormal ($\sigma=1, \mu = -3.5$)
Stretched exponential				
DDC	2.1	1.0	$0 \leq DDC \leq 0.1$	Lognormal ($\sigma=1, \mu = -3.5$)
α	0.67	0.8	$0 \leq \alpha \leq 1$	Uniform
Mono-exponential				
ADC	1.59	1.0	$0 \leq ADC \leq 0.1$	NA

D_p, D_t, DDC , and ADC are expressed in $\cdot 10^{-3} \text{ mm}^2/\text{s}$. f is expressed in %. NA = not applicable

Appendix B Parameter estimation errors

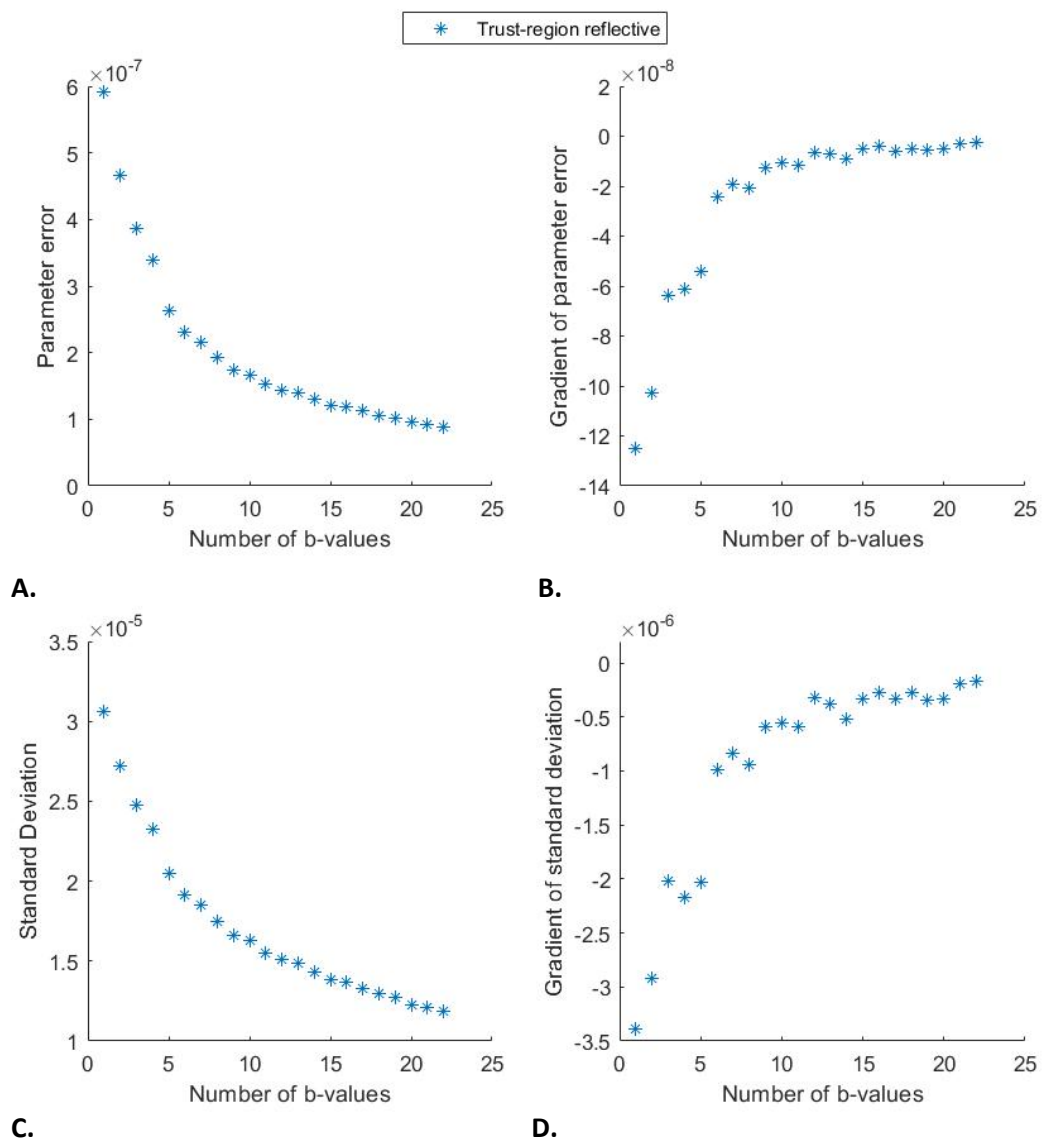


Figure B1. Results of fitting of mono-exponential apparent diffusion coefficient (ADC) per added b -value. A: error in parameter fitting, B: Gradient of fitting error at each set of b -values, C: Standard deviation of fitted parameters, D: Gradient of standard deviation at each set of b -values

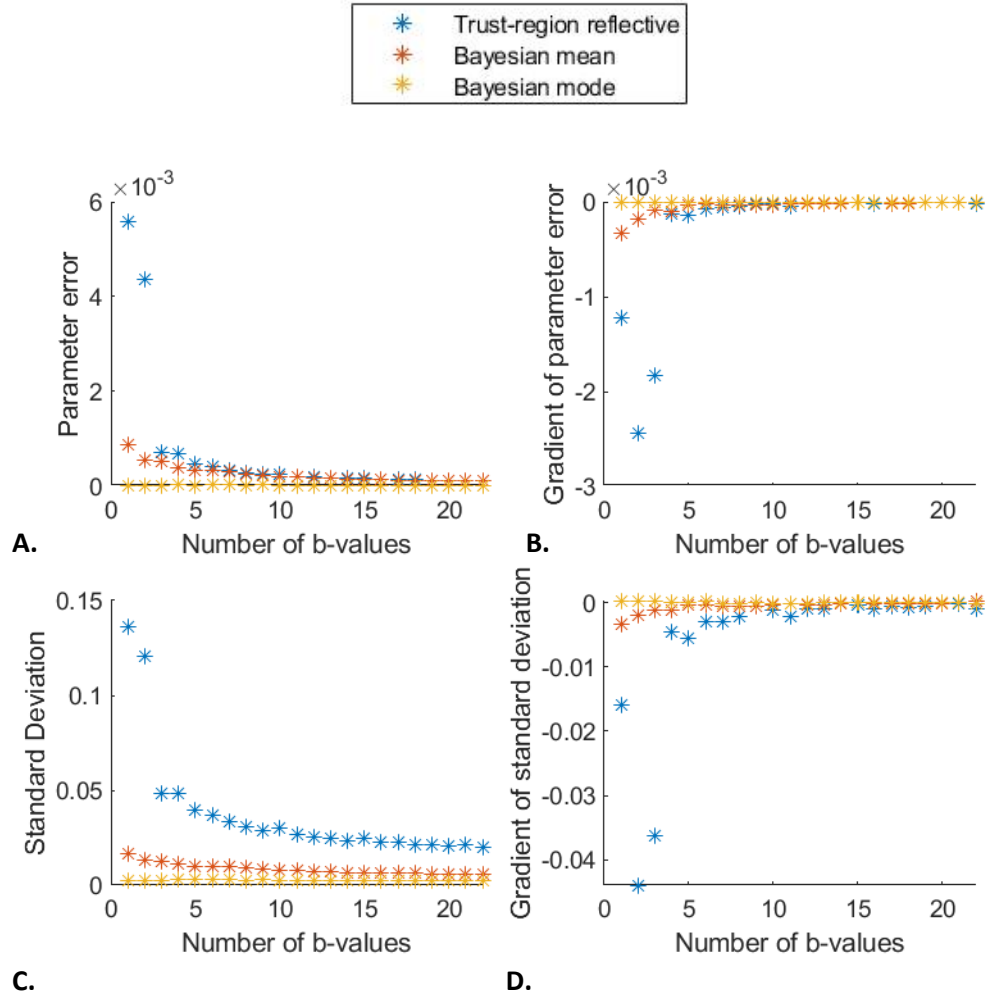


Figure B2 Results of fitting of intravoxel incoherent motion parameter f per added b -value. A: error in parameter fitting, B: Gradient of fitting error at each set of b -values, C: Standard deviation of fitted parameters, D: Gradient of standard deviation at each set of b -values

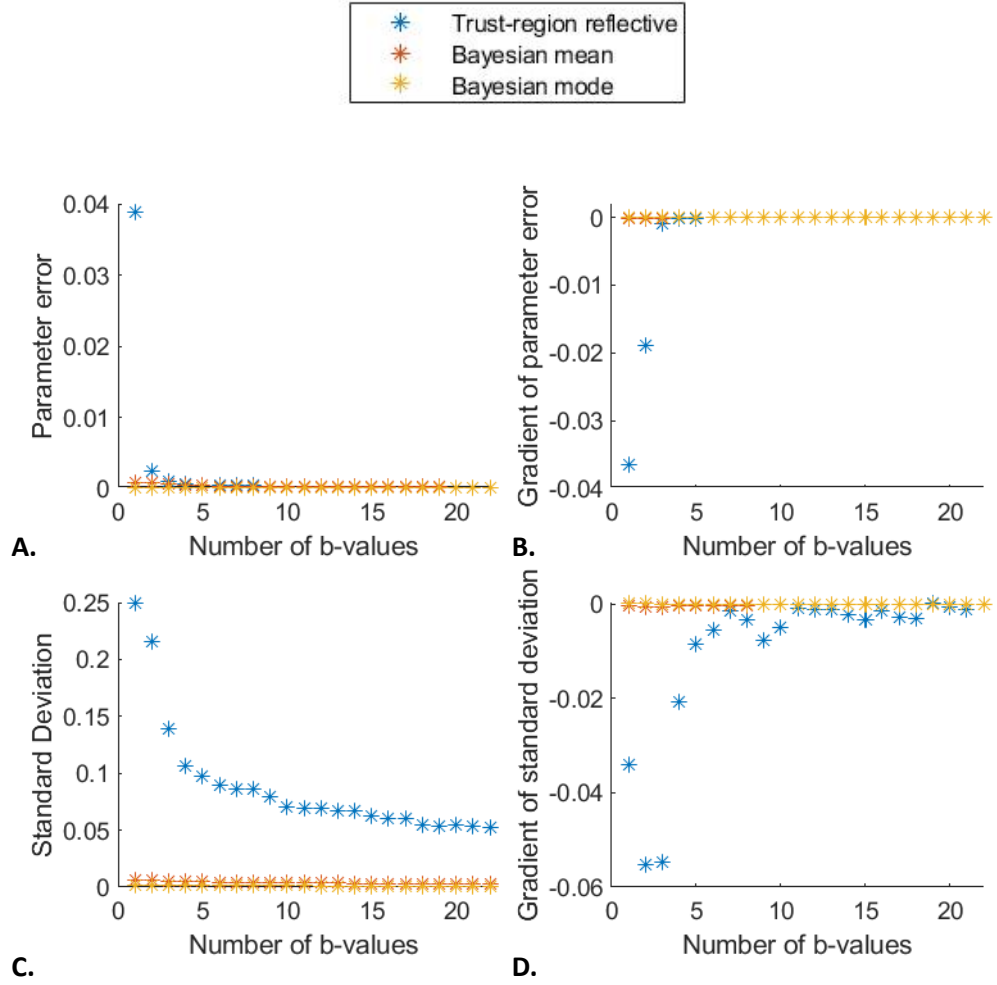


Figure B3. Results of fitting of intravoxel incoherent motion parameter D_p per added b -value. A: error in parameter fitting, B: Gradient of fitting error at each set of b -values, C: Standard deviation of fitted parameters, D: Gradient of standard deviation at each set of b -values

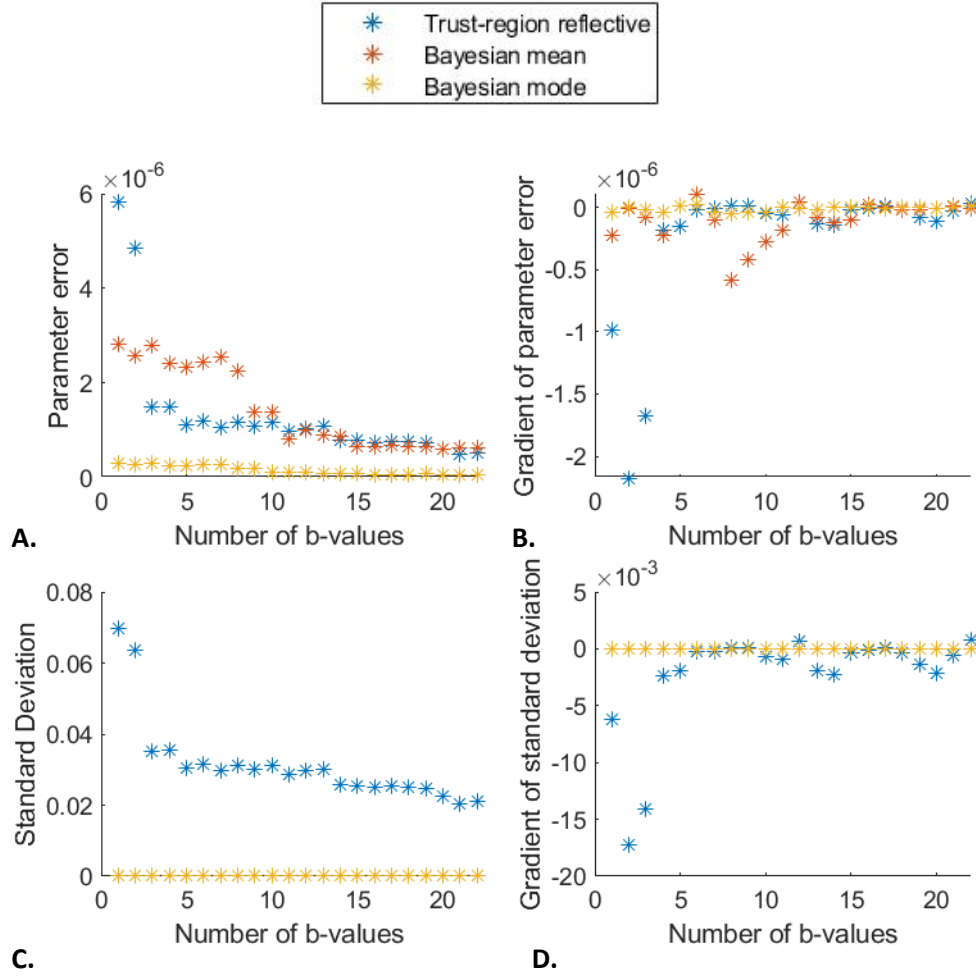


Figure B4. Results of fitting of intravoxel incoherent motion parameter D_t per added b -value. A: error in parameter fitting, B: Gradient of fitting error at each set of b -values, C: Standard deviation of fitted parameters, D: Gradient of standard deviation at each set of b -values

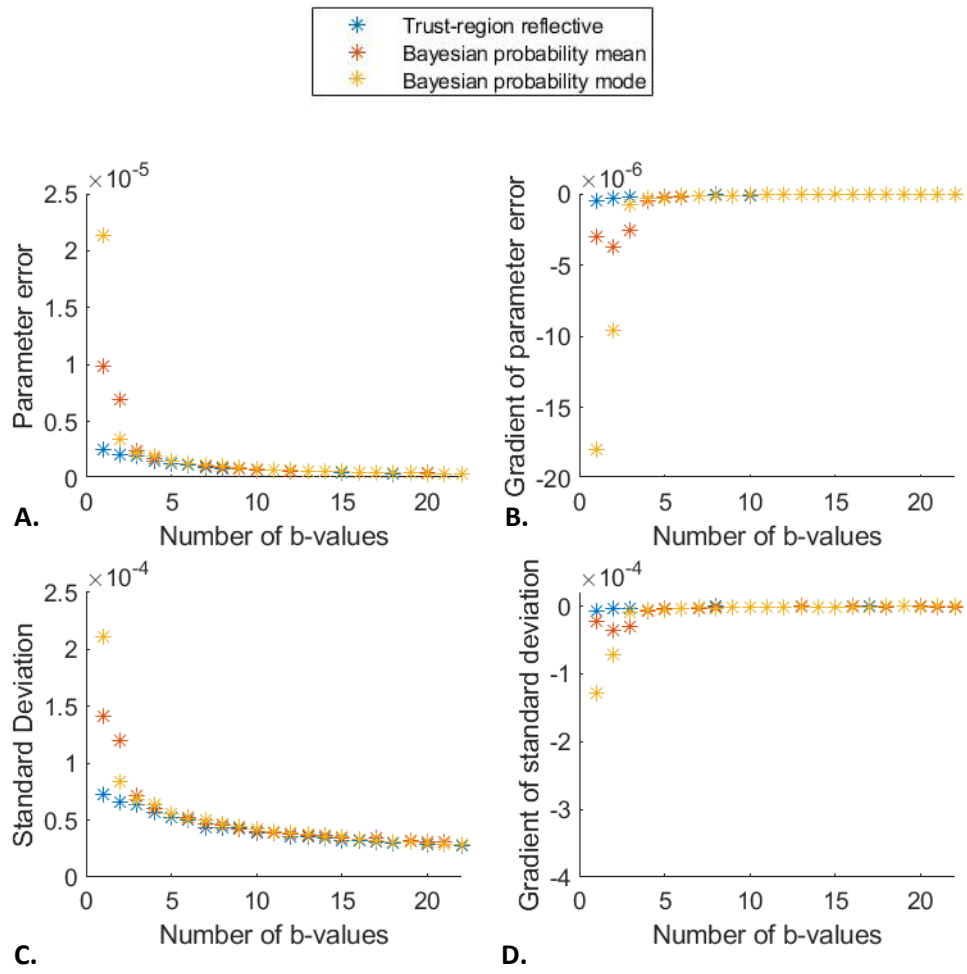


Figure B5. Results of fitting of stretched exponential parameter DDC per added b-value for each of the algorithms. A: error in parameter fitting, B: Gradient of fitting error at each set of b-values, C: Standard deviation of fitted parameters, D: Gradient of standard deviation at each set of b-values

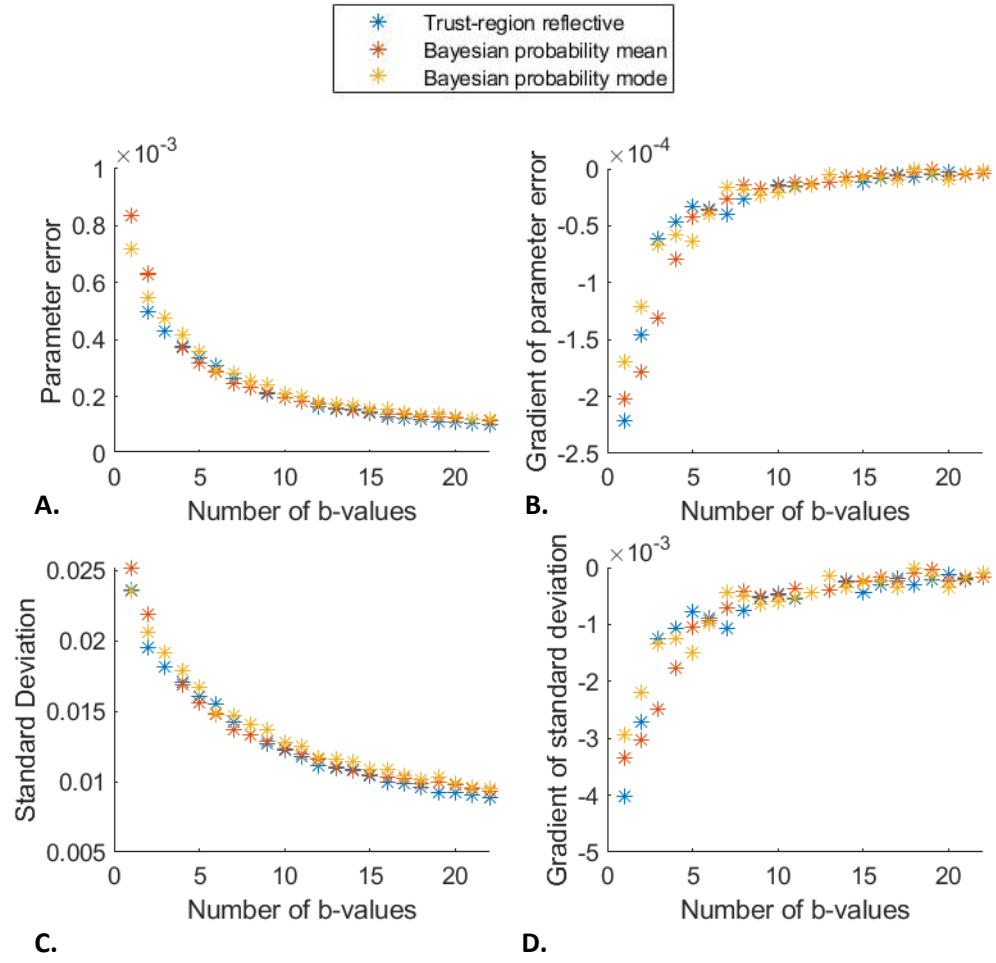


Figure B6. Fitting results of stretched exponential parameter α per number of b-values added to the initial b-value set for each of the algorithms. A: error in parameter fitting, B: Gradient of fitting error at each set of b-values, C: Standard deviation of fitted parameters, D: Gradient of standard deviation at each set of b-values

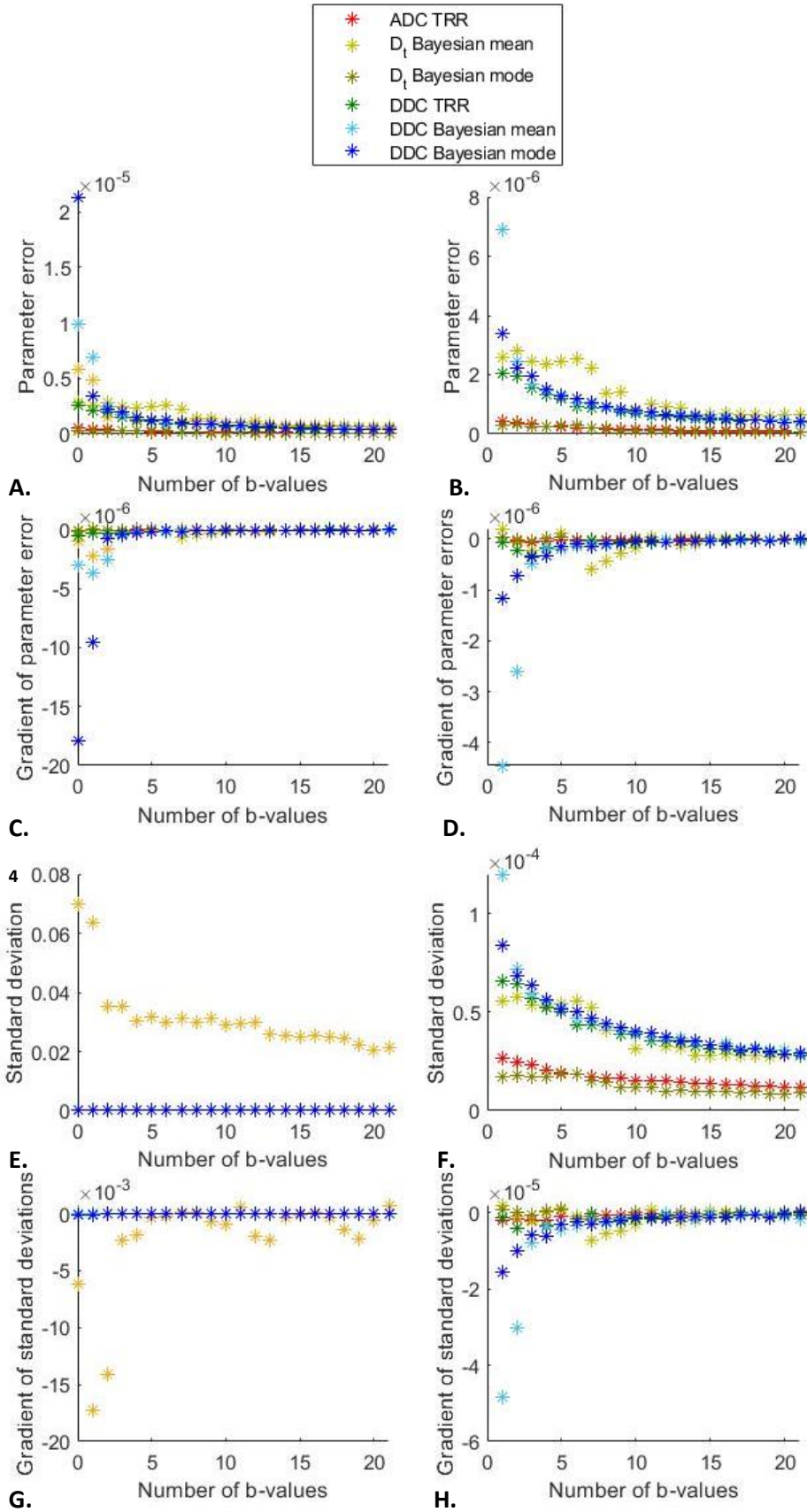


Figure B7. Fitting results of all diffusion parameters (D_t , DDC, and ADC) per number of b -values added to the initial b -value set, for each of the algorithms. Top row: errors in parameter fitting, Second row: Gradient of fitting errors, Third row: Standard deviations of fitted parameters, Bottom row: Gradient of standard deviations. The graphs in the right column show data obtained with five or more b -values, and data of the TRR algorithm for D_t are removed because of their large errors. ADC: apparent diffusion coefficient, D_t : true diffusion parameter, DDC: distributed diffusion coefficient, TRR: trust region reflective algorithm.

Appendix C Parameter errors and standard deviations for combinations of multi-exponential DWI-models

Table C1. DWI parameter estimation errors of the optimized *b*-value distribution

DWI model parameter	Estimation errors		
	TRR	BP mean	BP mode
IVIM			
f	$4.6 \cdot 10^{-4} \pm 3.9 \cdot 10^{-2}$	$3.42 \cdot 10^{-4} \pm 1.01 \cdot 10^{-2}$	$2.57 \cdot 10^{-5} \pm 2.76 \cdot 10^{-3}$
D_p	$4.76 \cdot 10^{-4} \pm 9.76 \cdot 10^{-2}$	$4.65 \cdot 10^{-4} \pm 4.82 \cdot 10^{-3}$	$4.55 \cdot 10^{-5} \pm 1.50 \cdot 10^{-3}$
D_t	$1.11 \cdot 10^{-6} \pm 3.04 \cdot 10^{-2}$	$2.34 \cdot 10^{-6} \pm 5.31 \cdot 10^{-5}$	$2.47 \cdot 10^{-7} \pm 1.71 \cdot 10^{-5}$
Stretched			
DDC	$1.31 \cdot 10^{-6} \pm 5.24 \cdot 10^{-5}$	$1.47 \cdot 10^{-6} \pm 5.55 \cdot 10^{-5}$	$1.51 \cdot 10^{-6} \pm 5.62 \cdot 10^{-5}$
α	$3.34 \cdot 10^{-4} \pm 1.60 \cdot 10^{-2}$	$3.18 \cdot 10^{-4} \pm 1.56 \cdot 10^{-2}$	$3.60 \cdot 10^{-4} \pm 1.67 \cdot 10^{-2}$

Data are presented as bias \pm variability. TRR: trust-region reflective, BP mean: Bayesian probability with mean as central tendency measure, BP mode: Bayesian probability with mode as central tendency

Table C2. Parameter errors and standard deviation for all combinations DWI-models and fitting algorithms

DWI model	Fitting algorithm					
IVIM	TRR	TRR	TRR	BPmean	BPmean	BPmean
Stretched	TRR	BPmean	BPmode	TRR	BPmean	BPmode
Error	0.0014 \pm 0.184	0.0013 \pm 0.183	0.0013 \pm 0.184	0.0012 \pm 0.031	0.0011 \pm 0.030	0.0012 \pm 0.032
DWI model	Fitting algorithm					
IVIM	BPmode	BPmode	BPmode			
Stretched	TRR	BPmean	BPmode			
Error	0.0004 \pm 0.021	0.0004 \pm 0.020	0.0004 \pm 0.022			

Data are presented as bias \pm variability. TRR: trust-region reflective, BP mean: Bayesian probability with mean as central tendency measure, BP mode: Bayesian probability with mode as central tendency

Appendix D Comparison of central tendency measures of the Bayesian probability algorithm in multi-exponential DWI-parameter histogram parameters

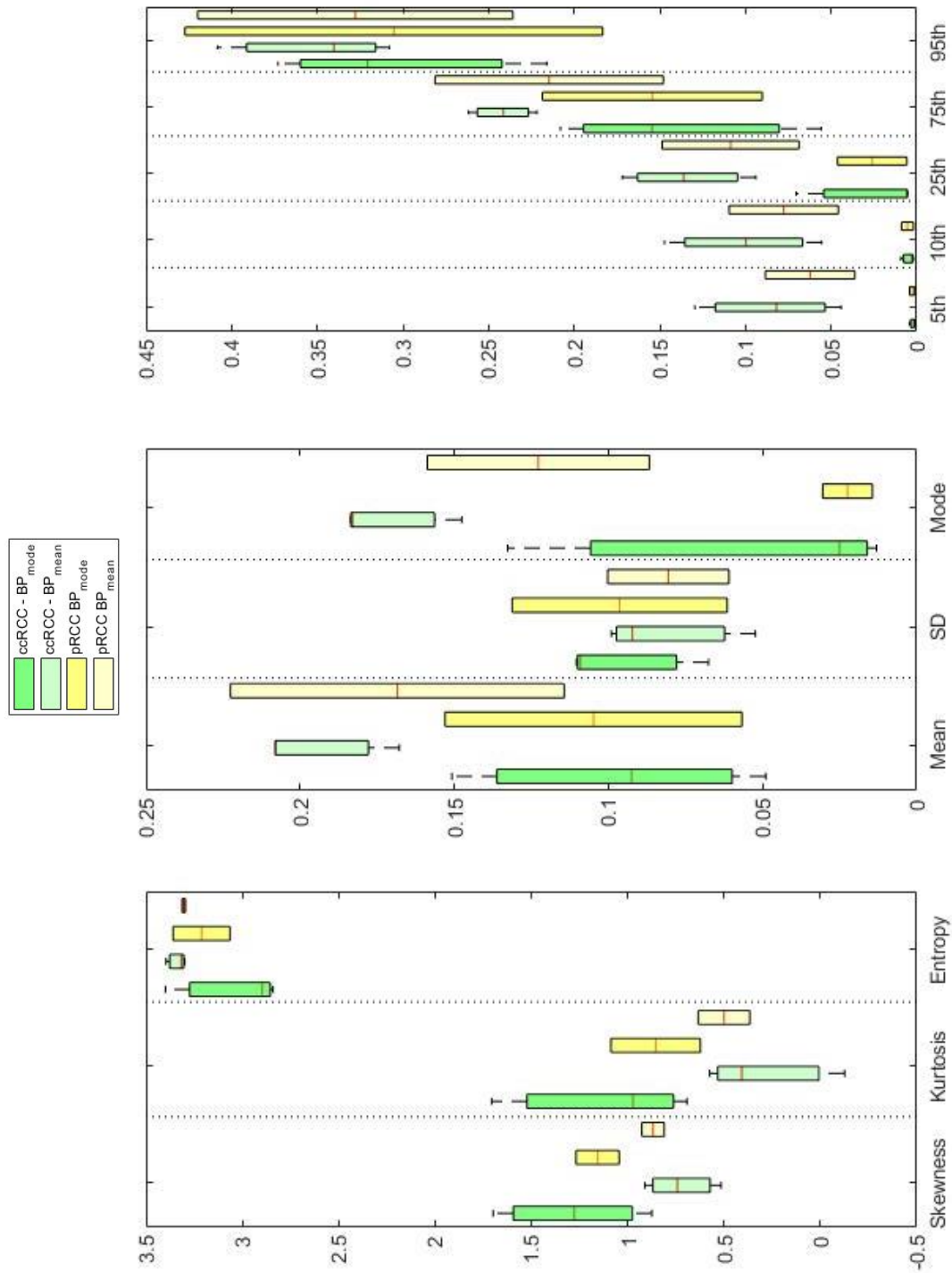


Fig D1. Intravoxel incoherent motion parameter f

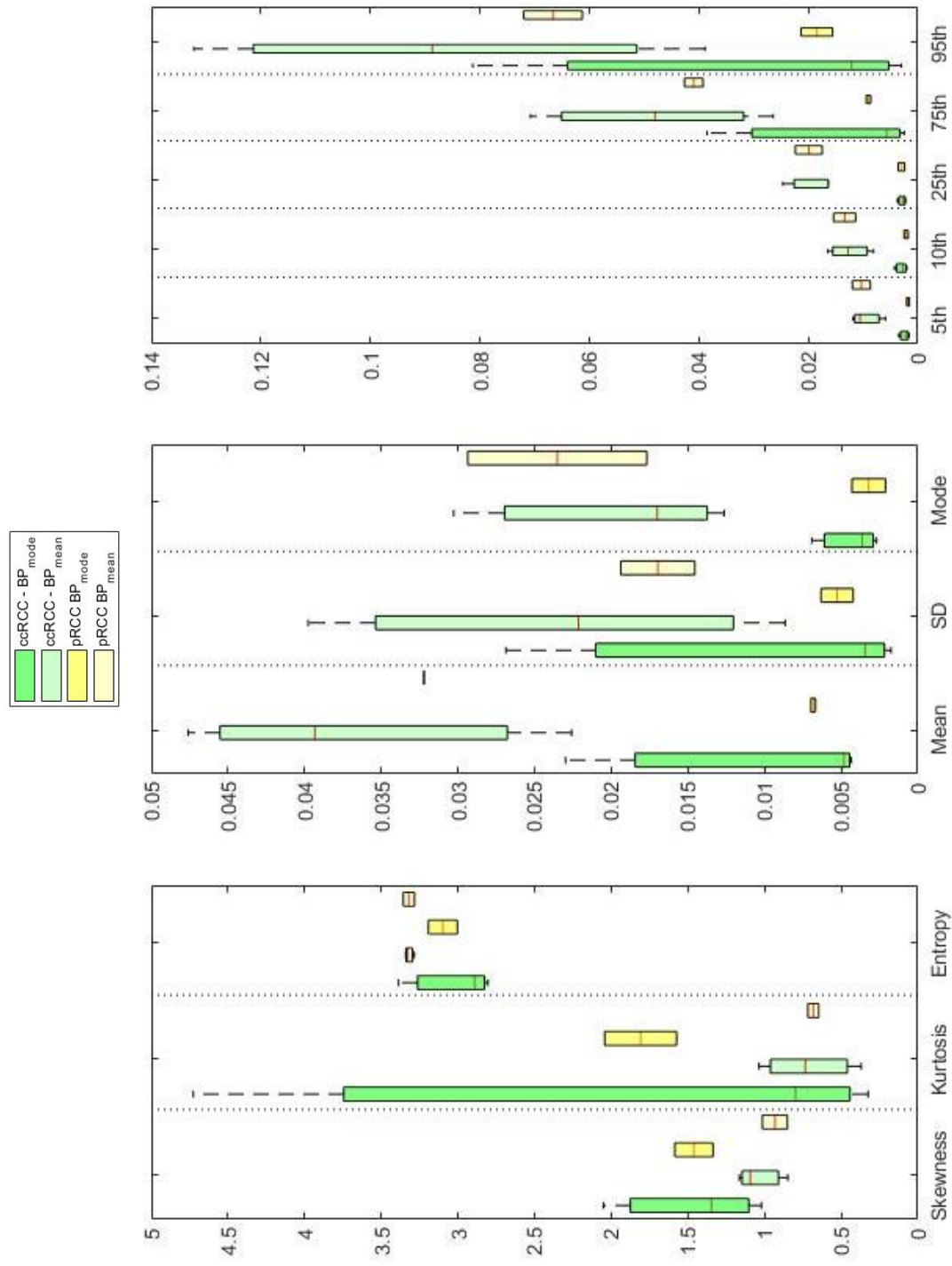


Fig D2. Intravoxel incoherent motion parameter D_p

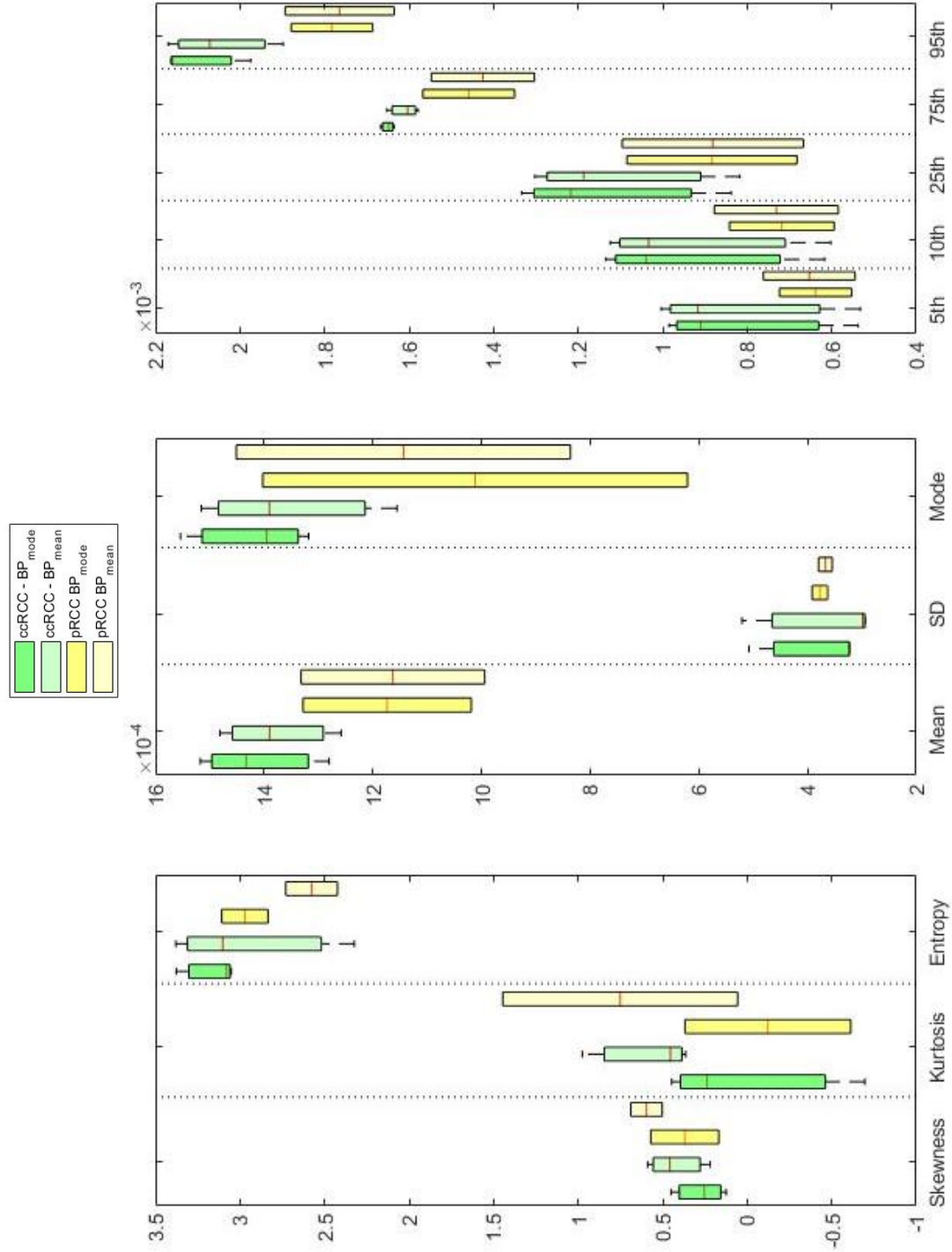


Fig D3. Intravoxel incoherent motion parameter D_t

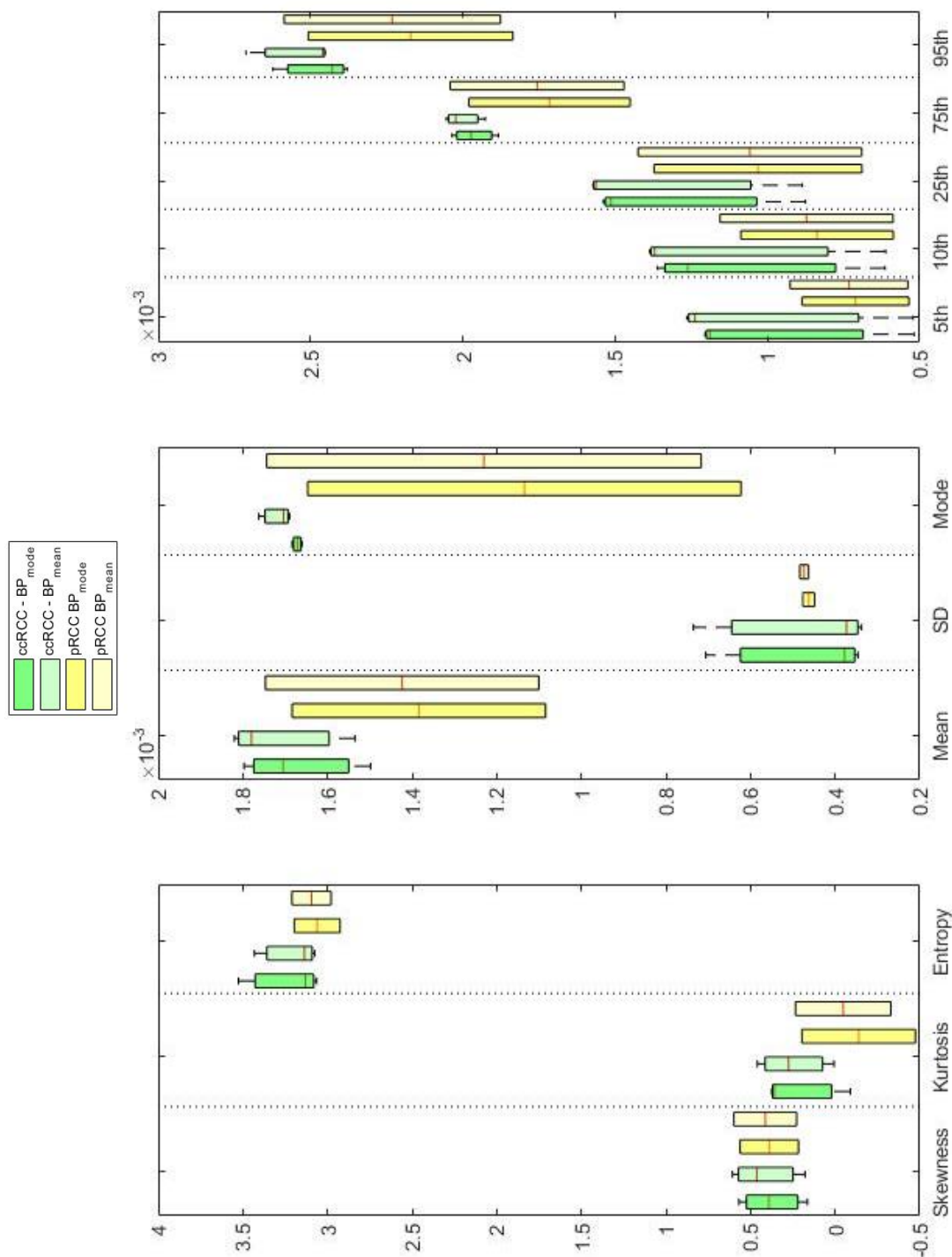


Fig D4. Stretched exponential parameter DDC

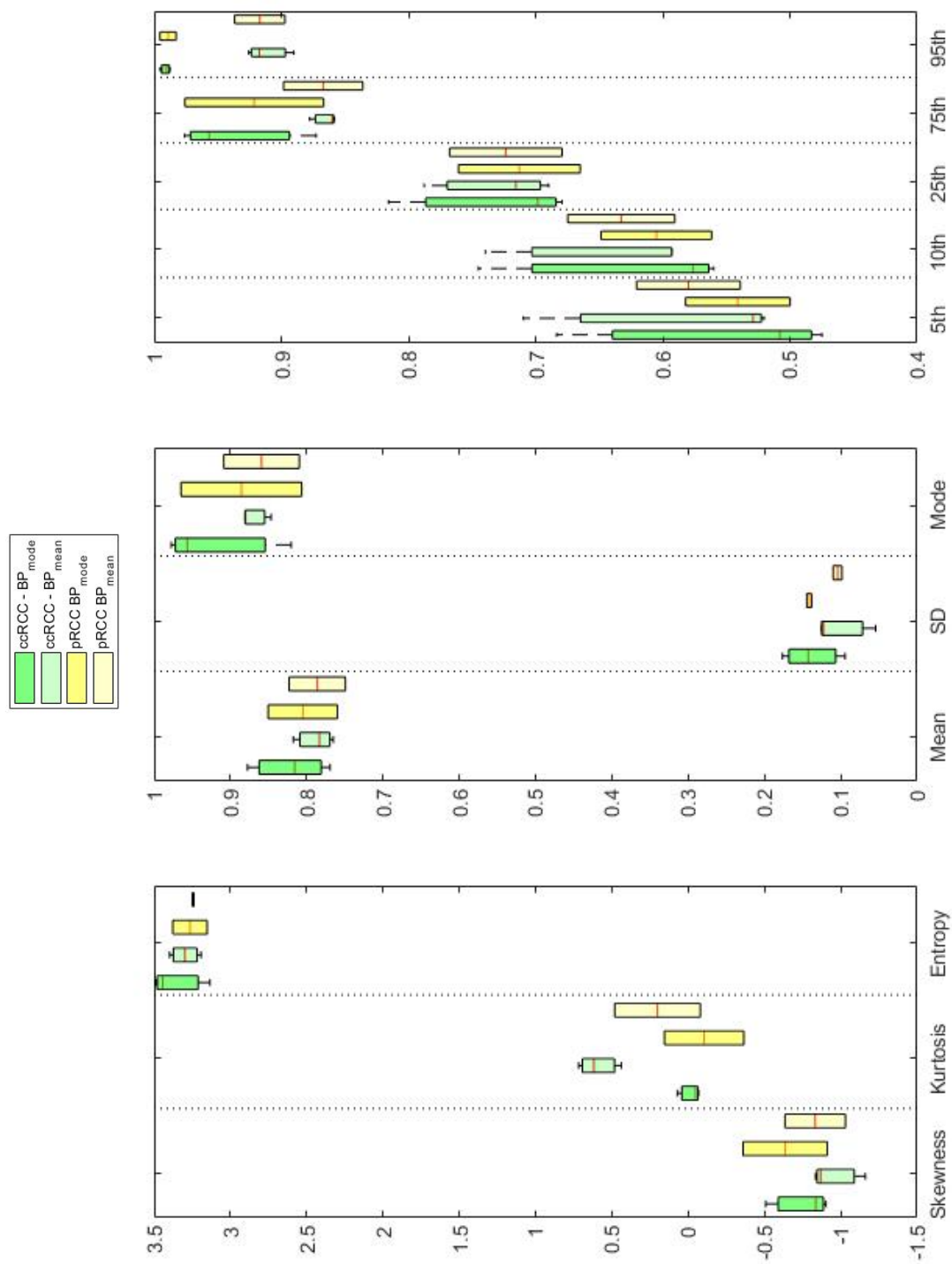
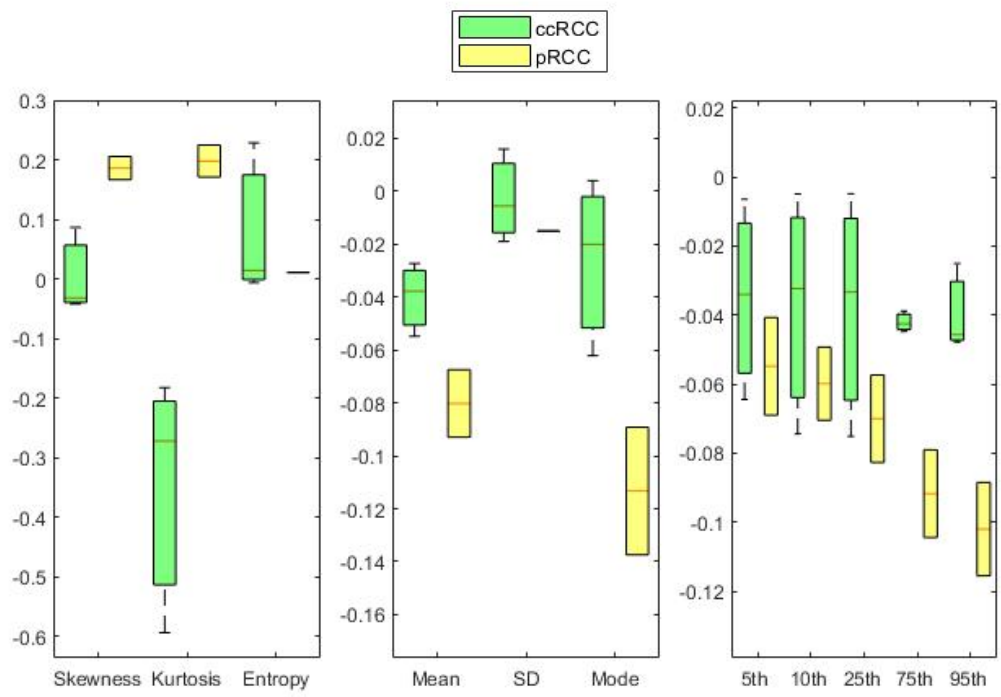
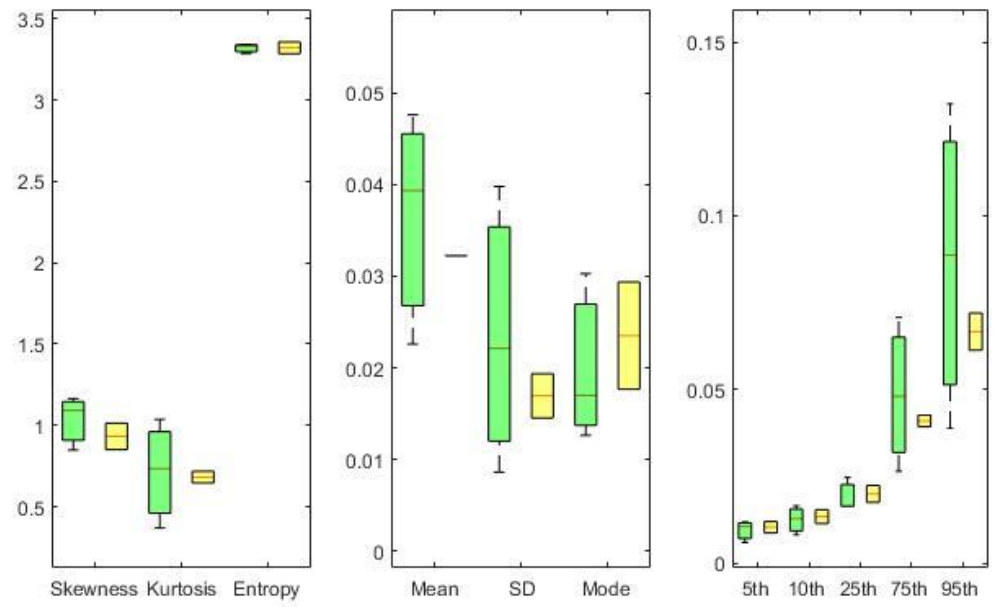


Fig D5. Stretched exponential parameter α

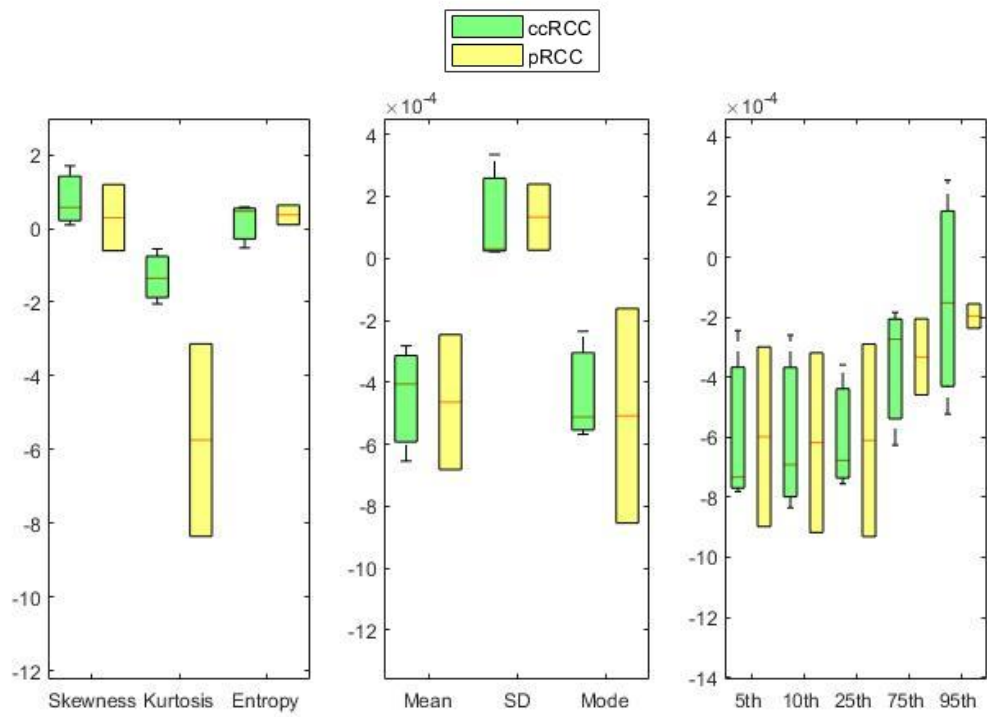
Appendix E Multi-exponential DWI-parameter histogram analysis of tumor induced differences



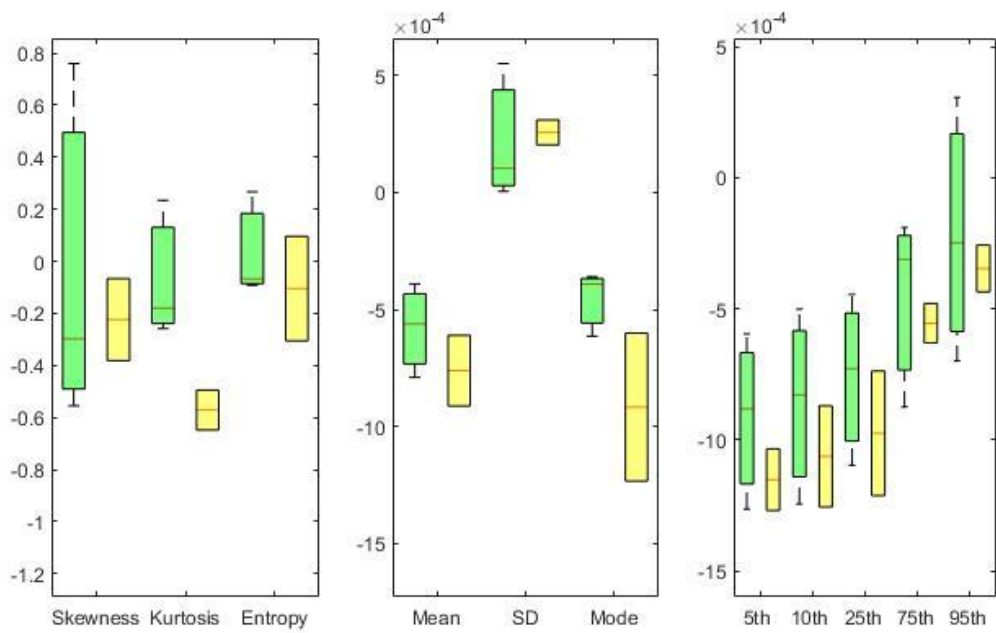
A.



B.



C.



D.

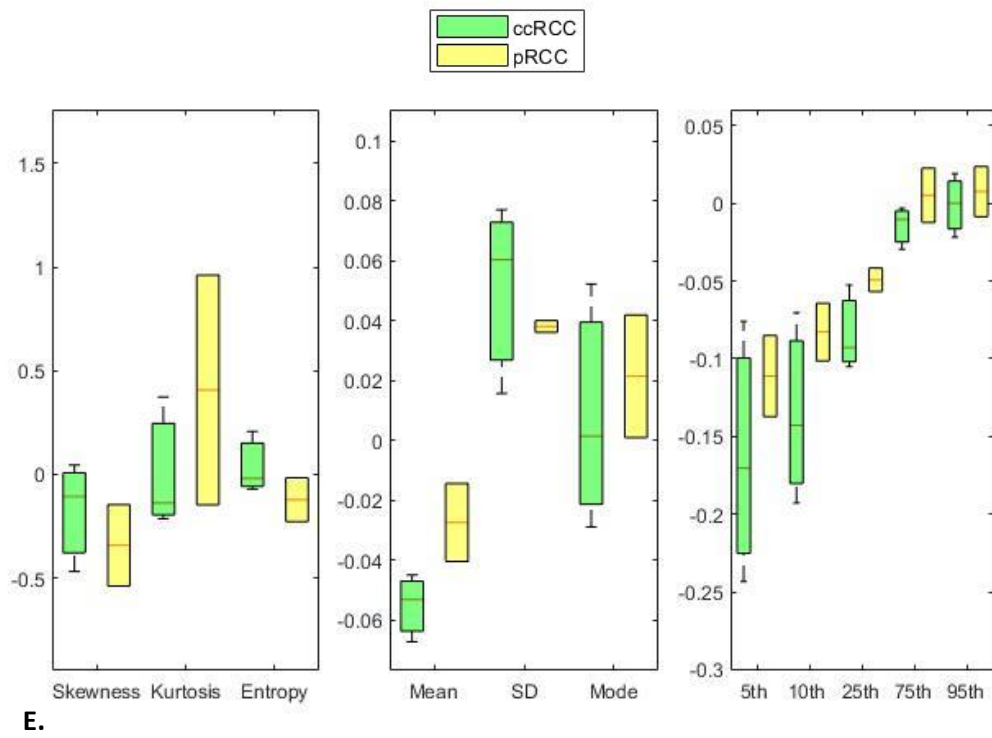


Figure E1. Histogram parameters of tumor induced differences (calculated by Tumor histogram parameter – contralateral cortical tissue parameter). Values for IVIM parameters f , D_p , and D_t are shown in (A), (B), and (C), respectively. Values for stretched exponential parameters DDC and α are shown in (D) and (E), respectively