

The Histogram Analysis of Diffusion-Weighted Intravoxel Incoherent Motion (IVIM) Imaging for Differentiating the Gleason grade of Prostate Cancer

Yu-Dong Zhang · Qing Wang · Chen-Jiang Wu ·
Xiao-Ning Wang · Jing Zhang · Hui Liu · Xi-Sheng Liu ·
Hai-Bin Shi

Received: 24 June 2014 / Revised: 2 October 2014 / Accepted: 14 November 2014 / Published online: 28 November 2014
© European Society of Radiology 2014

Abstract

Objective To evaluate histogram analysis of intravoxel incoherent motion (IVIM) for discriminating the Gleason grade of prostate cancer (PCa).

Methods A total of 48 patients pathologically confirmed as having clinically significant PCa (size >0.5 cm) underwent preoperative DW-MRI (b of 0–900 s/mm²). Data was post-processed by monoexponential and IVIM model for quantitation of apparent diffusion coefficients (ADCs), perfusion fraction f , diffusivity D and pseudo-diffusivity D^* . Histogram analysis was performed by outlining entire-tumour regions of interest (ROIs) from histological–radiological correlation. The ability of imaging indices to differentiate low-grade (LG,

Gleason score (GS) ≤6) from intermediate/high-grade (HG, GS >6) PCa was analysed by ROC regression.

Results Eleven patients had LG tumours (18 foci) and 37 patients had HG tumours (42 foci) on pathology examination. HG tumours had significantly lower ADCs and D in terms of mean, median, 10th and 75th percentiles, combined with higher histogram kurtosis and skewness for ADCs, D and f , than LG PCa ($p < 0.05$). Histogram D showed relatively higher correlations ($r = 0.641–0.668$ vs. ADCs: 0.544–0.574) with ordinal GS of PCa; and its mean, median and 10th percentile performed better than ADCs did in distinguishing LG from HG PCa.

Conclusion It is feasible to stratify the pathological grade of PCa by IVIM with histogram metrics. D performed better in distinguishing LG from HG tumour than conventional ADCs.

Key Points

- GS had relatively higher correlation with tumour D than ADCs.
- Difference of histogram D among two-grade tumours was statistically significant.
- D yielded better individual features in demonstrating tumour grade than ADC.
- D^* and f failed to determine tumour grade of PCa.

Y.-D. Zhang · Q. Wang · C.-J. Wu · X.-N. Wang · J. Zhang ·
X.-S. Liu · H.-B. Shi (✉)

Department of Radiology, the First Affiliated Hospital with Nanjing Medical University, 300, Guangzhou Road, Nanjing, Jiangsu Province, China 210009
e-mail: shihb@njmu.edu.cn

Y.-D. Zhang
e-mail: njmu_zyd@163.com

Q. Wang
e-mail: njmu_wq@163.com

C.-J. Wu
e-mail: njmu_wcj@163.com

X.-N. Wang
e-mail: njmu_wxn@163.com

J. Zhang
e-mail: njmu_zj@163.com

X.-S. Liu
e-mail: njmu_lxs@163.com

H. Liu
MR Collaboration NE Asia, Siemens Healthcare, Shanghai, China
e-mail: lh.liu@siemens.com

Keywords DWI · IVIM · Prostate cancer · Gleason score · MRI

Introduction

Pretreatment characterization of the tumour aggressiveness of prostate cancer (PCa) is vital for selecting the optimal therapy, and thus helpful to avoid overtreatment and to improve active surveillance for the patients with and without clinically significant cancers; however, exactly how the patients should be

risk stratified is still not clear [1–3]. Many nomograms and methods of stratifying PCa by risk classification have been reported for discriminating low-aggressiveness lesions from clinically significant PCa [4–6]. Of all the clinically determinable parameters, the Gleason scoring system has proven most important for measuring aggressiveness, disease outcome and the risk of mortality from PCa [3]. However, the estimation of Gleason score (GS) is based on invasive biopsy or radical prostatectomy (RP), which can be associated with serious adverse effects [7]. Therefore, development of a reliable non-invasive method that could differentiate low- from high-aggressiveness PCa would be a major advance and would have a significant benefit for individualized treatment options.

Recently, diffusion-weighted (DW) MR imaging combined with various mathematical models has been gaining substantial interest as a possible tool to detect localized PCa and predict its biochemical aggressiveness [8–14]. Documented studies have shown statistically significant correlations between the apparent diffusion coefficients (ADCs) obtained from DW imaging and the GS of PCa [15–18]. In addition, a few studies demonstrated that ADCs may predict prognosis for patients undergoing active surveillance and biochemical recurrence after RP [19, 20]. Although it is argued in most studies that the ADCs between different Gleason grade closely correlate with tumour cellularity, it is important to realize that monoexponential ADC measurement is partly influenced by tissue perfusion [21, 22]. The contribution of perfusion to the diffusion signal was elucidated by Le Bihan et al. in their pioneering work on intravoxel incoherent motion (IVIM) [23, 24]. The blood flow in a randomly oriented microvasculature, referred to as pseudo-diffusion, contributes to diffusion signal decay predominantly at low b values (<200 s/mm²). Perfusion effects can be resolved from the true tissue diffusion by acquiring DWI with a sufficiently wide range of b values, followed by biexponential curve fitting. Such an analysis can resolve pseudo-diffusivity D^* and tissue diffusivity D separately, along with their respective volume fractions f [25–27]. In recent years, several studies examined prostate DW-IVIM in the comparison of cancerous regions and normal tissue [22, 28, 29], but the confounding role of DW-IVIM for grading of tumour aggressiveness is rarely reported. In addition, the conventionally used method for quantitation of DW imaging data is limited by the performance with a region of interest (ROI)-based measurement. Such a method (e.g. placing regional ROIs on a representative section of the tumour) has been pointed out as a limitation of many studies [30, 31], in which the overlap of a single measurement may lead to interobserver variability in ROI placement. A histogram analysis approach has been shown to be a promising tool in discriminating tumour grade, differentiating their subtypes, and assessing therapeutic effects on the cancer [32–34].

The purpose of our study was thus to primarily assess the diagnostic performance of IVIM for stratifying the aggressiveness of PCa on the basis of an entire-tumour histogram analysis, by using post-RP pathological results as reference standard.

Materials and methods

Patients

This retrospective study was approved by our institutional review board and written informed consent was waived. Between December 2012 and March 2014, 63 consecutive patients with clinically localized prostate cancer (proved by biopsy examination) who were referred for routine clinical evaluation underwent standard MR examination (including DW imaging) before RP. The inclusion criteria were (a) patient had undergone no prior hormonal or radiation treatment, (b) DW imaging performed with the same parameters, (c) the diagnosis of “primary” PCa and (d) performance of detailed histopathology where at least one tumour focus with a diameter of 0.5 cm or more available for multiparametric DW-MR calculation. A total of 48 patients met all inclusion criteria and were included in this study. The other patients were excluded because sequence parameters of DW imaging were not unified ($n=7$), the tumour was too small ($n=5$) or neoadjuvant therapy was administered before the MR examination ($n=3$).

MR acquisition

MR examinations were performed with a 3.0-T MR system (Verio Tim; Siemens, Erlangen, Germany) and a pelvic phased-array coil. As per the standard clinical prostate MR examination at our institution, the images obtained included transverse T₁-weighted turbo spin-echo (TSE) images (repetition time (ms)/echo time (ms), 700/14; section thickness, 3.5 mm; intersection gap, 0.5 mm; field of view, 25 cm; and matrix, 384×336) and transverse, coronal and sagittal T₂-weighted TSE images (repetition time (ms)/effective echo time (ms) 6,000/124; section thickness, 3.5 mm; intersection gap, 0.3 mm; field of view, 25 cm; and matrix, 384×336) of the prostate and seminal vesicles. Then, single-shot echo-planar imaging (repetition time (ms)/echo time (ms), 6,000/72; field of view, 25 cm; matrix, 192×130; section thickness, 3.5 mm; intersection gap, 0.3 mm; a parallel imaging factor of 2; and 21 sections) was performed with diffusion-module and fat suppression pulses. Diffusion in three directions was measured by using b values of 0, 50, 150, 300, 600 and 900 s/mm².

Histological–radiological correlation

After RP, the prostatic specimens were uniformly processed and submitted for histological investigation. The prostatectomy specimens were fixed in 10 % neutral buffered formalin and stored overnight after surgical resection. Prostatectomy specimens were fixed in 5 % buffered formalin, processed and cut serially into 4-mm-thick blocks from apex to base in transverse planes. Each block was then halved or quartered (depending on its size), and 7–8- μm -thick microtome slices were stained with haematoxylin and eosin. A genitourinary pathologist with more than 10 years' experience in genitourinary pathology reviewed all the sections and outlined the location of the tumour on the photographs for each slice. Tumours were graded according to the 2005 International Society of Urological Pathology Modified Gleason Grading System.

We referred to the studies by Peng et al. [35] and Oto et al. [36] for the histological–radiological correlation. During this procedure, the pathologist identified all distinct tumour foci larger than 0.5 cm in diameter, and a radiologist (Y.Z., 8 years of experience in urinogenital imaging) manually outlined the corresponding regions of interest (ROIs) of the tumour foci on axial T2-weighted MR images. For those cancer foci that were not clearly visible on MR images, their locations were determined on the basis of their relationship with other identifiable landmarks (e.g. urethra, ejaculatory ducts, benign prostatic hyperplasia nodules) on MR images by consensus of the radiologist and pathologist. The drawn ROIs on MR images were required to carefully match the extent of the tumour determined from histological examination slice by slice. Gleason score was reported specifically for each tumour ROI by the study pathologist during the consensus review.

MR post-processing

All data were transferred in Digital Imaging and Communications in Medicine (DICOM) format and processed off-line using MATLAB software (MathWorks Inc. Natick, MA, USA). The DW data were post-processed with the monoexponential and biexponential IVIM model, respectively. For the biexponential IVIM model, the relationship between signal intensity of DWI and b factors can be expressed by Eq. 1:

$$\frac{S(b)}{S_0} = f e^{-D^*b} + (1-f)e^{-D} \quad (1)$$

where $S(b)$ is the mean signal intensity, S_0 is the signal intensity without diffusion, f is the fraction of perfusion, influenced by directional flow of water molecules during diffusion time. D is the diffusion parameter representing pure molecular diffusion (the slow component of diffusion), and

D^* is the diffusion parameter representing incoherent micro-circulation within the voxel, i.e. perfusion-related diffusion or the fast component of diffusion. Data were fitted with the Levenberg–Marquardt nonlinear least squares algorithm. Pixel-by-pixel maps of diffusion features, D , D^* , f and ADC were then automatically constructed from the proposed models.

In terms of ROI measurement, the obtained ROIs on T2-weighted MR images were transferred to all other multiparametric DW-IVIM imaging maps with computer software developed in-house, which allowed one the manual adjustment of the potential misalignment. Lastly, the multiparametric measurement of ROIs in each slice of the tumour foci was summated to derive voxel-by-voxel values for a histogram analysis. The histogram of DW-IVIM parameters for each ROI was analysed by using commercially available software (PASW Statistics 16.0; SPSS Inc., Chicago, IL, USA) for histogram analysis. D , D^* and ADC histograms were plotted with diffusivity on the x-axis with a bin size of $1 \times 10^{-6} \text{ mm}^2/\text{s}$, and the f histogram was plotted with perfusion fraction on the x-axis with a bin size of 1×10^{-3} . The y-axis expressed the percentage of tumour volume by dividing the frequency in each bin by the total number of voxels analysed. For further quantitative analysis, cumulative D , D^* , f and ADC values were obtained from their histograms, respectively, in which the cumulative number of observations in all of the bins up to the specified bin was mapped onto the y-axis as a percentage. Based on an entire-tumour measurement, the following histogram parameters were derived from ADC, D , D^* and f maps, respectively: (a) mean; (b) median; (c) kurtosis, which is the degree of peakedness of a distribution and (d) skewness, which is a measure of the degree of asymmetry of a distribution. For the cumulative histograms, the 10th and 75th percentiles of the tumour D , D^* , f and ADCs were derived, respectively (the n th percentile is the point at which n % of the voxel values that form the histogram are found to the left). A representative case for the introduction of metrics for histological–radiological correlation and histogram analysis of DW imaging measures are shown in Fig. 1.

Statistical analysis

All statistical analyses were performed by using SPSS 16.0. P values of less than 0.05 determined statistical significance. We firstly calculated the Spearman rank-order correlation coefficient (\tilde{r}) to characterize the strength of correlation between multiparametric imaging measures and ordinal GS. And as discriminating the low-grade (LG) from combined intermediate- and high-grade (HG) tumours using the quantitative DW imaging measures is of clinical importance, we classified the data into LG group ($\text{GS} \leq 6$) and HG group ($\text{GS} > 6$), respectively. The normality and homoscedasticity of the multiparametric imaging data were tested using the

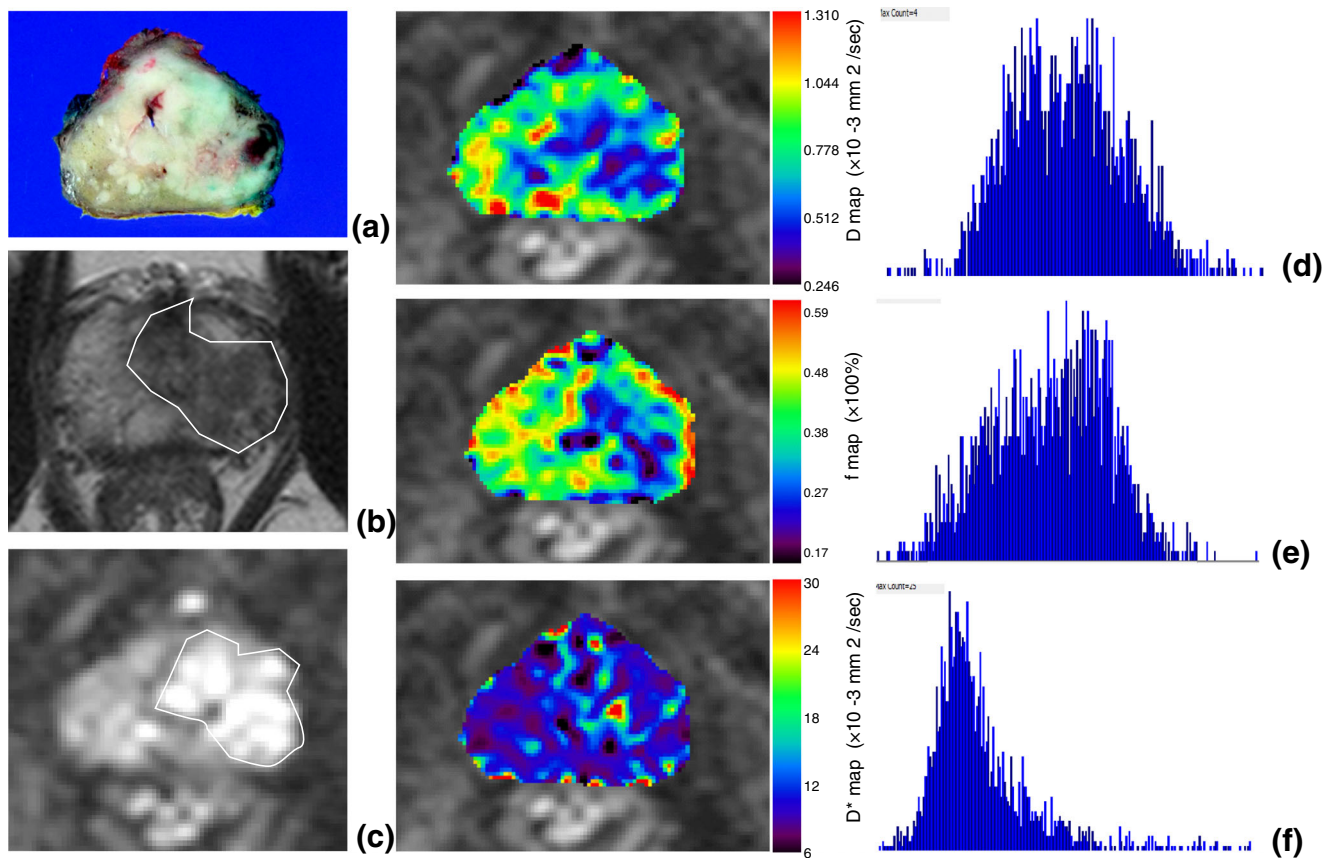


Fig. 1 A representative PCa with Gleason score of 4+4 for the introduction of metrics for histologic–radiologic correlation, and histogram analysis of DW imaging measures. A solid mass was shown in the left peripheral zone from a histologic specimen (a), and characteristic appearance with inhomogeneous hypointense signal

intensity (SI) on axial T2WI (b) and hyperintense SI on DWI (c). The tumour boundary was identified by histologic–radiologic correlation and was then outlined on T2WI and DWI (white line). The pixel-by-pixel D (d), f (e) and D^* (f) were obtained, respectively, and the corresponding histogram distributions were constructed

Q–Q plots and Levene’s tests. Data satisfying the assumption (mean, median, the 10th and 75th percentiles) were subjected to independent sample t test. Conversely, data not satisfying the assumption (kurtosis, skewness) were analysed by using the Mann–Whitney U test. The overall ability to discriminate between LG and HG tumour was analysed by using a receiver operating characteristic (ROC) regression model, and quantified by using the areas under the ROC curves (A_z), referring to the method of DeLong et al. [37]. The diagnostic accuracy, sensitivity and specificity were calculated at a cutoff point that maximized the value of the Youden index. In the case of the 15 out of 48 patients who had more than one tumour focus, making the analysis not independent of each other, only the tumour focus with the highest GS was considered for ROC analysis.

Results

Of the total of 48 histologically confirmed cases, 30 had only peripheral zone (PZ) PCa. In the remaining 18 patients, five

had only transitional zone (TZ) PCa and 13 had both PZ and TZ PCa. The median serum PSA level for the cohort was 19.2 ng/mL (range, 0.7–214.4 ng/mL; median, 19.2 ng/mL). Of these 48 patients, 33 patients had one tumour focus, five patients had two tumour foci (one had a diameter of tumour foci less than 0.5 cm), two patients had three tumour foci, eight patients had more than four tumour foci (six patients had only one tumour foci with a diameter larger than 0.5 cm, two patients had three tumour foci with a diameter larger than 0.5 cm). After exclusion of the tumour foci with a diameter less than 0.5 cm, a total of 60 tumour foci were included for the end imaging measures. The median volume of the included tumour foci was 2.3 ml (0.5–61.5 ml). Clinical characteristics of patients, as well as those of tumour foci ROIs, are summarized in Table 1.

Spearman’s rank correlation showed that all the histogram indices for D reflected statistically significant correlations with ordinal GS ($p < 0.01$). And for histogram ADCs, only mean, median and 10th and 75th percentiles reflected statistically significant correlations with ordinal GS ($p < 0.01$). All the histogram indices for D^* and f did not reflect statistically significant correlations with ordinal GS ($p > 0.05$). It was noted

Table 1 Summary of clinical and pathologic characteristics

Variable	Value
Clinical characteristics*	
Age (years)	70 (57–86)
PSA (ng/ml)	19.2 (0.7–214.4)
Tumour size (ml)	2.3 (0.5–61.5)
Pathologic characteristics†	
pT2a	9 (18.7)
pT2b	11 (22.9)
pT2c	7 (14.6)
pT3a	11 (22.9)
pT3b	6 (12.5)
pT4	4 (8.3)
Gleason grade‡	
GS 3+3	11 (18)
GS 3+4	8 (13)
GS 4+3	11 (11)
GS 4+4	8 (8)
GS 4+5	6 (6)
GS 5+3	1 (1)
GS 5+4	3 (3)

* Data are medians, with ranges in parentheses

† Data are no. of patients, with percentages in parentheses

‡ Data are no. of patients, with no. of tumour foci in parentheses

that the indices for D exhibited relatively higher correlation coefficients with tumour GS than those of ADCs (Table 2).

The independent sample t test showed that the HG tumour had significantly lower ADCs than LG tumour in terms of histogram mean, median and the 10th and 75th percentiles (all $p < 0.001$). For IVIM-derived D , the histogram mean, median and the 10th and 75th percentiles reflected statistically significant differences between LG and HG groups (all $p < 0.001$). For D^* and f , the mean, median and 10th and 75th percentiles for D^* and f did not reflect statistically significant differences between the two qualitative groups (all $p > 0.05$). The Mann–Whitney U test showed that, for histogram ADCs, D and f , the HG tumours had both higher kurtosis and skewness than LG tumours (all $p < 0.010$). The histogram kurtosis and skewness for D^* did not reflect statistically significant differences between the two qualitative groups (Table 3). Two representative cases with LG and HG PCa for the comparison of

Table 2 Spearman's rank correlation of tumour Gleason grade to DW-IVIM histograms

Histogram variable	ADC	D	D^*	f
Mean	-0.553†	-0.662†	0.069	-0.023
Median	-0.574†	-0.668†	0.057	-0.031
10th percentile	-0.555†	-0.659†	-0.081	0.025
75th percentile	-0.544†	-0.641†	0.070	-0.010
Kurtosis	0.132	0.384†	-0.035	0.214
Skewness	0.158	0.359†	-0.122	0.176

† Correlation is significant at the 0.01 level (2-tailed)

multiparametric DW imaging measures are shown in Figs. 2 and 3.

Table 4 demonstrates the results of the ROC analyses of the multiparametric imaging measures for distinguishing patients with LG tumour from patients with HG tumour. The difference of A_z was not significant between each histogram index for ADCs and f (all $p > 0.05$). For IVIM-derived D , the difference of A_z was significant between kurtosis vs. mean ($p = 0.017$), vs. median ($p = 0.016$) and vs. 10th ($p = 0.031$), respectively. And the difference of A_z was significant between skewness vs. mean ($p = 0.017$), vs. median ($p = 0.016$) and vs. 10th percentile ($p = 0.009$), respectively. In terms of the comparison between D and ADCs, D had significantly higher A_z values than ADCs in histogram mean ($p = 0.018$), median ($p = 0.044$) and 10th percentile ($p = 0.023$), respectively. And these histogram indices derived from D exhibited relatively higher accuracy, sensitivity and specificity for discrimination between LG and HG tumours than did ADCs (Fig. 4).

Discussion

Our study demonstrates that stratification of the Gleason grade of PCa is feasible by multiparametric DW-IVIM MR imaging with histogram analysis metrics. The patients with PCa were determined as those with a dominant decrease in tumour diffusivity (reflected by ADCs or D), corresponding to the increased Gleason grade. And interestingly, the HG tumour had higher kurtosis and skewness for histogram ADCs, D and f than LG tumour, indicating more heterogeneity and complexity of cellularity in a tumour with a high Gleason score. Moreover, the histogram mean, median and 10th percentile D exhibited better individual features for discriminating patients with LG tumour from patients with HG tumour than did ADCs, indicating that the proposed IVIM method is a more precise way of stratifying biotic features of PCa. However, the difference of histogram D^* and f between the two qualitative grade groups did not reach statistical significance, which indicates that pseudo-perfusion may contribute little to the diffusivity for predicting the tumour grade of PCa.

DW-MRI has been accepted as an important imaging biomarker of PCa and provides reliable indices for grading of various tumour entities [16, 18, 38]. In recent years, several studies examined IVIM in patients with PCa, but inconsistent results were presented (e.g. for f , both higher and lower values compared to normal tissue were reported) [22, 39, 40]; and moreover, the behaviour of IVIM parameters in these studies was only compared between normal tissue and biopsy findings, and the confounding role of DW-IVIM for grading of tumour aggressiveness is rarely reported. In the present study, we employed the IVIM analysis to evaluate the aggressiveness of tumour foci in two qualitative grade groups, and compared these results to conventional ADCs. It

Table 3 Cumulative histogram parameters of DW-IVIM according to Gleason score

Variable	Low grade ($n=18$)	Intermediate/high grade ($n=42$)	<i>P</i> value
Histogram ADCs ($\times 10^{-6}$ mm ² /s)			
Mean	1,497 \pm 208	1,272 \pm 255	0.002 [†]
Median	1,441 \pm 211	1,183 \pm 234	<0.001 [†]
10th percentile	1,125 \pm 235	867 \pm 222	<0.001 [†]
75th percentile	1,655 \pm 204	1,413 \pm 254	0.001 [†]
Kurtosis [‡]	1.8 (21.8)	10.7 (34.2)	0.012 [†]
Skewness [‡]	1.19 (21.6)	2.54 (34.3)	0.010 [†]
Histogram <i>D</i> ($\times 10^{-6}$ mm ² /s)			
Mean	1,172 \pm 150	934 \pm 181	<0.001 [†]
Median	1,159 \pm 158	904 \pm 195	<0.001 [†]
10th percentile	928 \pm 163	669 \pm 137	<0.001 [†]
75th percentile	1,303 \pm 166	1,063 \pm 203	<0.001 [†]
Kurtosis [‡]	-0.18 (19.9)	1.19 (35.1)	0.002 [†]
Skewness [‡]	0.38 (22.7)	0.80 (33.8)	0.023 [†]
Histogram <i>D</i> * ($\times 10^{-6}$ mm ² /s)			
Mean	16,277 \pm 2,739	16,944 \pm 1,963	0.290
Median	14,333 \pm 3,171	14,704 \pm 3,574	0.705
10th percentile	7,192 \pm 2,371	7,059 \pm 2,322	0.841
75th percentile	23,878 \pm 5,696	24,410 \pm 4,387	0.696
Kurtosis [‡]	-1.09 (31.8)	-1.10 (29.9)	0.705
Skewness [‡]	0.54 (33.4)	0.40 (29.2)	0.392
Histogram <i>f</i> (%)			
Mean	12.1 \pm 3.4	13.3 \pm 5.1	0.362
Median	11.2 \pm 3.1	12.1 \pm 5.2	0.487
10th percentile	4.8 \pm 2.9	4.7 \pm 2.3	0.917
75th percentile	15.6 \pm 3.8	17.1 \pm 7.7	0.443
Kurtosis [‡]	1.06 (21.3)	3.41 (34.4)	0.008 [†]
Skewness [‡]	0.82 (21.9)	1.33 (34.2)	0.012 [†]

Unless otherwise specified, data are means \pm standard deviations

[‡]Data are medians, with mean ranks (obtained by Mann–Whitney *U* test) in parentheses

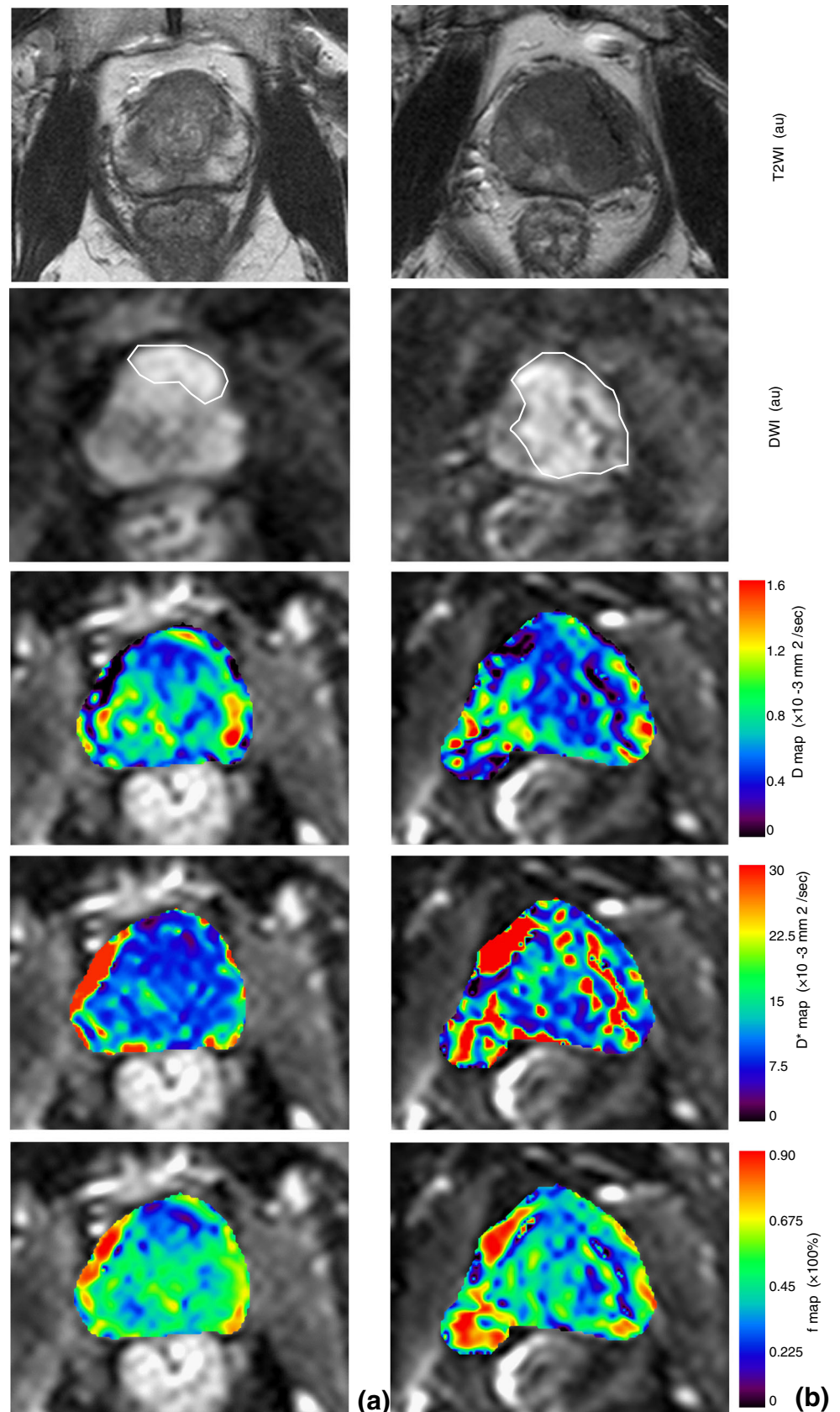
[†]Difference is significant at the given level

demonstrated that parameter *D* is excellent for discriminating tumour Gleason grade, and has better individual features (higher *Az* values) for discriminating low-grade tumour foci from intermediate/high-grade tumour foci. The predominant decrease in *D* within HG tumour is commonly understood to be a consequence of increased tumour cellularity, which results in a more decreased diffusivity of water molecules [41–43]. In addition, within the prostate, the contribution of DW imaging signal is also from extracellular components, e.g. from tubular structures, luminal space and their fluid content [44, 45]. Therefore, the biexponential fitting model derived from IVIM is theoretically more favourable for describing the tumour biological characteristics. A statistical difference of *f* has been previously indicated between cancerous zones and healthy individuals by Quentin and Pang et al. [21, 22]. However, there was no significant difference of both *f* and *D** in two Gleason-grade groups in our study, which indicated that the extracellular components, such

as fluid content or vascular perfusion, made little contribution to DW imaging signal in the cancerous zone, and were thus little related to its aggressiveness. This was supported by a similar study [36], which demonstrated that quantitative DCE-MRI parameters did not show a significant correlation with Gleason score or VEGF expression, and limited studies evaluating the correlation between DCE-MRI parameters and angiogenesis markers of PCa have also provided contradicting results [16, 46, 47].

It is well known that increased heterogeneity and complexity of tumour cellularity, one of the important histological features in malignant prostate cancer tissues, is highly responsible for the local aggressiveness and pathological grade of the tumour [48, 49]. Within the PCa, tissue heterogeneity could produce a wide spectrum of cellular density in a localized tumour site. This means that ADCs within the whole tumour volume may vary widely between different regions of the tumour. Unfortunately, the commonly used method for DW

Fig. 2 DW-IVIM imaging of a tumour with Gleason score (GS) of 3+3 (a) and a tumour foci with GS of 4+4 (b) from two different patients. It shows a heterogeneous low SI on T₂-weighted images and high SI on corresponding DW images ($b=600$ s/mm²) in both lesions. The *D* maps show the remarkably decreased diffusivity both in low- and high-GS tumour foci, but the high-GS tumour shows a relatively low *D* compared to low-GS tumour. The tumour in both cases show a slightly decreased perfusion fraction compared to normal tissue. The *D** maps do not show evident difference between normal and cancerous regions both in low- and high-grade tumour foci



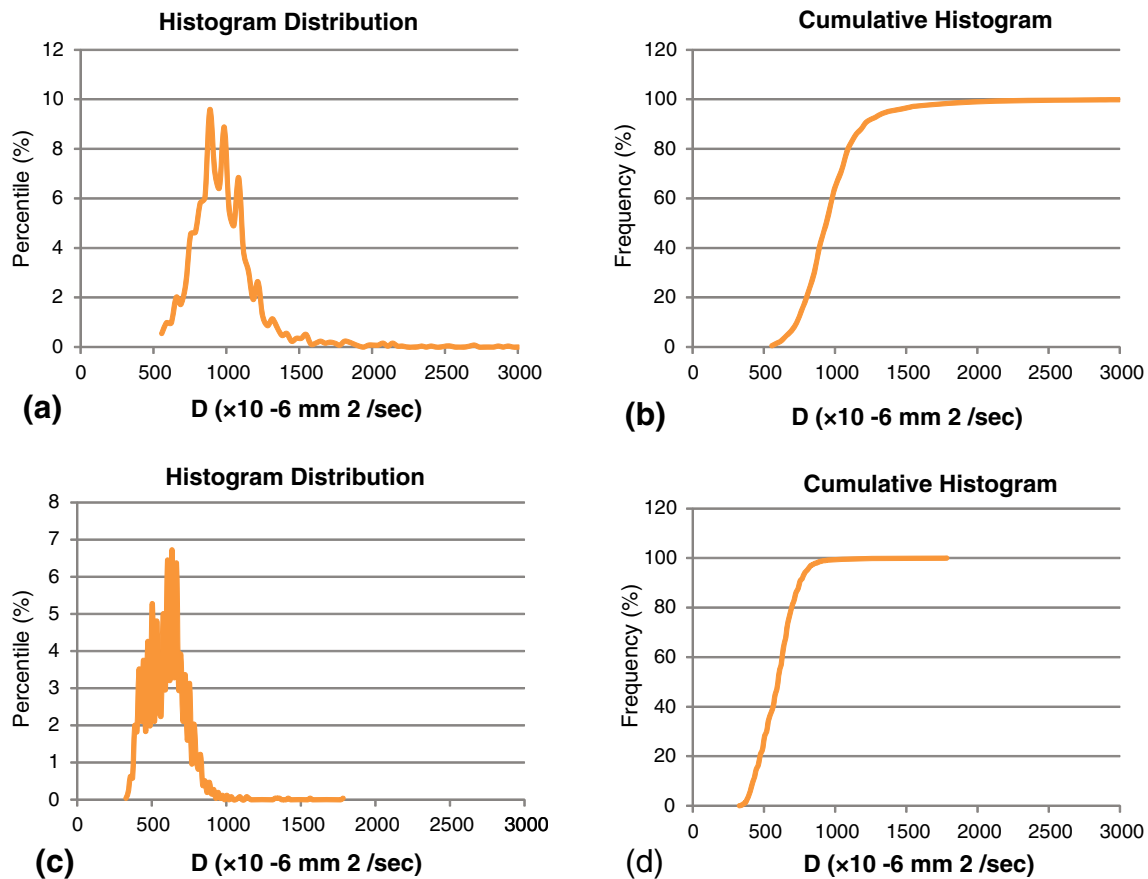


Fig. 3 Comparison of entire-tumour *D* histograms for two cases with GS 3+4 (a, b) and GS 5+4 (c, d). The tumour with GS 3+4 shows a lower relative frequency at low *D* compared with GS 5+4 tumour foci, resulting

in substantial divergence between low- and high-grade PCa at the low end of the cumulative histograms. This suggests that the high-grade PCa contained more pixels with low *D*, which indicates high cellularity

imaging measurement (e.g. placement of regional ROIs on a representative section of the tumour) has been pointed out as a limitation of many studies, in which the overlap of single ADC measurements between different grades did not accurately reflect the true features of tumour cellular density. We explored a histogram analysis for focusing more on the distribution of DW-IVIM parameters, rather than simple summary statistics. This histogram analysis approach was proved to be a promising tool in discriminating tumour grade or monitoring the effects of chemotherapy in brain and ovarian cancer as it is able to objectively reflect different microenvironments of diffusivity through the entire tumour volume [32, 34, 50, 51]. Interestingly, we illustrated its potential in discriminating tumour Gleason grade. The kurtosis and skewness, respectively, reflect the histogram distribution with regards to peakedness and normality; and the increase in kurtosis and skewness in HG PCa may represent more complexity of intratumorous cellularity that are characteristic of various Gleason-scored nuclei, necrosis, haemorrhage or fibrosis. The low-percentile histogram *D* may represent the microstructural information of tumour foci with the highest cellular density. Visually, the diffusivity changes translated into a shift of the histograms toward the left

and adoption of a more asymmetrical shape in tumour with HG PCa, indicating increasing nuclei density and higher risk of aggressiveness.

By using ROC curves, we found that our histogram *D* was able to distinguish tumours with intermediate to high Gleason grade from those with low Gleason grade with an accuracy of 72.9–91.7 %, sensitivity of 70.3–91.8 % and specificity of 63.6–90.9 %, which were relatively higher than those obtained with our histogram ADCs and conventional ADCs obtained by Rosenkrantz et al. [52]. The better diagnostic performance of the proposed IVIM method for discriminating tumour grade in our study was probably due to (a) an improvement in imaging measures, which was based on an entire-tumour histogram analysis and carefully matching histological-radiologic findings; and (b) the usage of biexponential curve fitting metrics, which was reported to show better individual features in terms of model fit and repeatability [11]. In addition, the histogram mean, median and 10th percentile *D* produced higher Az values than kurtosis and skewness, indicating a better diagnostic effectiveness of these indices for stratification of tumour grade. However the inter-group difference of Az between mean, median,

Table 4 Effectiveness of quantitative DW imaging indices in differentiating low-grade from intermediate/high-grade prostate cancer

Variable	Az ^a	Accuracy	Sensitivity	Specificity	Cutoff value ^b
Histogram ADCs					
Mean	0.811	66.7 (32/48)	59.4 (22/37)	90.9 (10/11)	≤1,311
Median	0.851	70.8 (33/48)	62.2 (23/37)	90.9 (10/11)	≤1,294
10th percentile	0.846	81.2 (39/48)	83.8 (31/37)	72.7 (8/11)	≤1,101
75th percentile	0.809	75.0 (36/48)	75.7 (28/37)	72.7 (8/11)	≤1,587
Kurtosis	0.761	81.2 (39/48)	83.8 (31/37)	72.7 (8/11)	>3.16
Skewness	0.750	83.3 (40/48)	91.8 (34/37)	54.5 (6/11)	>0.99
Histogram D^{\dagger}					
Mean [‡]	0.924	87.5 (42/48)	86.5 (32/37)	90.9 (10/11)	≤1,118
Median [‡]	0.924	91.7 (44/48)	91.8 (34/37)	90.9 (10/11)	≤1,101
10th percentile [‡]	0.955	87.5 (42/48)	86.5 (32/37)	90.9 (10/11)	≤839
75th percentile	0.897	87.5 (42/48)	89.2 (33/37)	81.8 (9/11)	≤1,282
Kurtosis	0.794	79.2 (38/48)	83.8 (31/37)	63.6 (7/11)	>0.11
Skewness	0.754	72.9 (35/48)	70.3 (26/37)	81.8 (9/11)	>0.54
Histogram f					
Kurtosis	0.750	79.2 (38/48)	83.8 (31/37)	63.6 (7/11)	> 1.19
Skewness	0.741	79.2 (38/48)	83.8 (31/37)	64.6 (7/11)	> 0.86

Unless otherwise specified, data are percentage, with raw data in parentheses

[†] Difference of Az is significant between kurtosis vs. mean ($p=0.017$), vs. median ($p=0.016$) and vs. 10th ($p=0.031$), respectively; and between skewness vs. mean ($p=0.017$), vs. median ($p=0.016$) and vs. 10th ($p=0.009$), respectively

[‡] Az is significantly higher than ADCs for histogram mean ($p=0.018$), median ($p=0.044$) and 10th ($p=0.023$)

^a Data are arbitrary units (au)

^b Units of $\times 10^{-6}$ mm²/s for mean, median and 10th and 75th percentiles; units of au for kurtosis and skewness

the 10th and 75th percentiles of D did not reflect statistical significance, indicating their similar diagnostic ability; moreover, the difference of histogram D^* and f between two qualitative groups did not reach statistical significance, demonstrating that pseudo-perfusion may contribute little to the diffusivity for predicting the tumour grade of PCa.

Our study also has several limitations. First, similar to the study of Peng et al. [35], we performed the histological–radiologic correlation through a systematic consensus-seeking correlative review of histological findings and MR

images by a genitourinary pathologist and a radiologist. Although this was carried out carefully to reduce the potential mismatch, uncertainty still cannot be ruled out in some cases. Second, it is a retrospective study with a relatively small sample size that could have been influenced by selection biases; the number of tumours with GS ≤ 6 was relatively small. This may produce statistical uncertainty with regards to relatively high true negative cases (high specificity) in the study. Therefore, larger prospectively studied patient populations will be needed to refine the correlation of IVIM parameters and tumour aggressiveness.

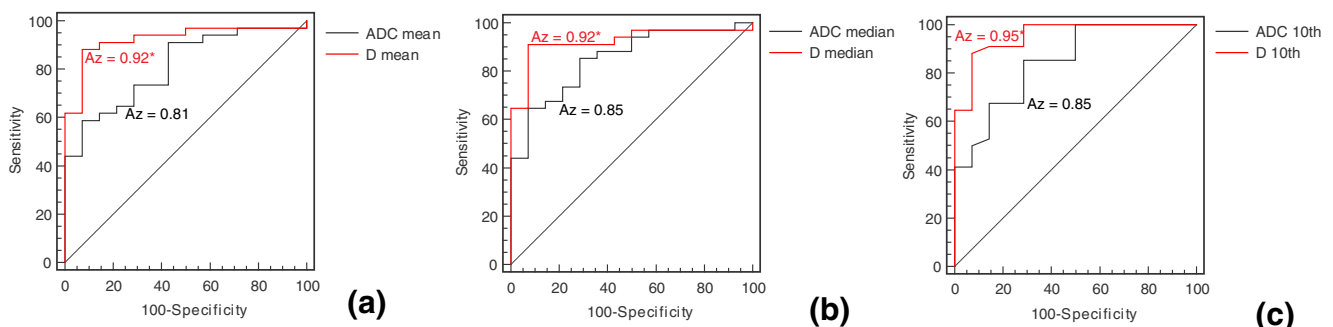


Fig. 4 Comparison of diagnostic ability for discriminating LG from HG tumour between histogram ADCs and D . The ROC analysis shows that D has significantly higher Az values than ADCs in histogram mean (a),

median (b) and 10th percentile (c), indicating a more reliable with IVIM method for stratifying tumour Gleason grade. *Difference is significant at the 0.05 level

Conclusion

Our results suggest that diffusivity D derived from IVIM can be a useful tool for discriminating low-grade tumour foci from intermediate/high-grade tumour foci in patients with PCa; and the f and D^* contribute little to the diffusivity for predicting the tumour grade. The histogram analysis is helpful to reflect the varieties of biologic behaviour in tumour foci.

Acknowledgments The scientific guarantor of this publication is Hai-Bin Shi, M.D. The authors of this manuscript declare no relationships with any companies whose products or services may be related to the subject matter of the article. The authors state that this work has not received any funding. One of the authors has significant statistical expertise. No complex statistical methods were necessary for this paper. Institutional review board approval was obtained. Written informed consent was waived by the institutional review board.

Methodology: retrospective, case-control study/diagnostic or prognostic study, performed at one institution

References

- Ruzsics B, Suranyi P, Kiss P et al (2008) Head-to-head comparison between delayed enhancement and percent infarct mapping for assessment of myocardial infarct size in a canine model. *J Magn Reson Imaging* 28:1386–1392
- Kattan M (2002) Statistical prediction models, artificial neural networks, and the sophism "I am a patient, not a statistic". *J Clin Oncol* 20:885–887
- D'Amico AV, Whittington R, Malkowicz SB et al (1998) Biochemical outcome after radical prostatectomy, external beam radiation therapy, or interstitial radiation therapy for clinically localized prostate cancer. *JAMA* 280:969–974
- Cooperberg MR, Pasta DJ, Elkin EP et al (2005) The University of California, San Francisco Cancer of the Prostate Risk Assessment score: a straightforward and reliable preoperative predictor of disease recurrence after radical prostatectomy. *J Urol* 173:1938–1942
- Stephenson AJ, Scardino PT, Eastham JA et al (2006) Preoperative nomogram predicting the 10-year probability of prostate cancer recurrence after radical prostatectomy. *J Natl Cancer Inst* 98:715–717
- Poulakis V, Witzsch U, de Vries R et al (2004) Preoperative neural network using combined magnetic resonance imaging variables, prostate specific antigen, and Gleason score to predict prostate cancer recurrence after radical prostatectomy. *Eur Urol* 46:571–578
- Namiki S, Saito S, Ishidoya S et al (2005) Adverse effect of radical prostatectomy on nocturia and voiding frequency symptoms. *Urology* 66:147–151
- Barentsz JO, Richenberg J, Clements R et al (2012) ESUR prostate MR guidelines 2012. *Eur Radiol* 22:746–757
- deSouza NM, Riches SF, Vanas NJ et al (2008) Diffusion-weighted magnetic resonance imaging: a potential non-invasive marker of tumour aggressiveness in localized prostate cancer. *Clin Radiol* 63:774–782
- Jung SI, Donati OF, Vargas HA, Goldman D, Hricak H, Akin O (2013) Transition zone prostate cancer: incremental value of diffusion-weighted endorectal MR imaging in tumor detection and assessment of aggressiveness. *Radiology* 269:493–503
- Jambor I, Merisaari H, Taimen P et al (2014) Evaluation of different mathematical models for diffusion-weighted imaging of normal prostate and prostate cancer using high b-values: a repeatability study. *Magn Reson Med*. doi:10.1002/mrm.25323
- Merisaari H, Jambor I (2014) Optimization of b-value distribution for four mathematical models of prostate cancer diffusion-weighted imaging using b values up to 2000 s/mm: simulation and repeatability study. *Magn Reson Med*. doi:10.1002/mrm.25310
- Tamura C, Shinmoto H, Soga S et al (2014) Diffusion kurtosis imaging study of prostate cancer: preliminary findings. *J Magn Reson Imaging* 40:723–729
- Quentin M, Pentang G, Schimmoller L et al (2014) Feasibility of diffusional kurtosis tensor imaging in prostate MRI for the assessment of prostate cancer: preliminary results. *Magn Reson Imaging* 32:880–885
- Vargas HA, Akin O, Franiel T et al (2011) Diffusion-weighted endorectal MR imaging at 3 T for prostate cancer: tumor detection and assessment of aggressiveness. *Radiology* 259:775–784
- Langer DL, van der Kwast TH, Evans AJ et al (2010) Prostate tissue composition and MR measurements: investigating the relationships between ADC, T2, K(trans), v(e), and corresponding histologic features. *Radiology* 255:485–494
- Turkbey B, Shah VP, Pang Y et al (2011) Is apparent diffusion coefficient associated with clinical risk scores for prostate cancers that are visible on 3-T MR images? *Radiology* 258:488–495
- Hambrock T, Somford DM, Huisman HJ et al (2011) Relationship between apparent diffusion coefficients at 3.0-T MR imaging and Gleason grade in peripheral zone prostate cancer. *Radiology* 259(2):453–461
- Morgan VA, Riches SF, Thomas K et al (2011) Diffusion-weighted magnetic resonance imaging for monitoring prostate cancer progression in patients managed by active surveillance. *Br J Radiol* 84:31–37
- Park SY, Kim CK, Park BK, Lee HM, Lee KS (2011) Prediction of biochemical recurrence following radical prostatectomy in men with prostate cancer by diffusion-weighted magnetic resonance imaging: initial results. *Eur Radiol* 21:1111–1118
- Quentin M, Blondin D, Klases J et al (2012) Comparison of different mathematical models of diffusion-weighted prostate MR imaging. *Magn Reson Imaging* 30:1468–1474
- Pang Y, Turkbey B, Bernardo M et al (2013) Intravoxel incoherent motion MR imaging for prostate cancer: an evaluation of perfusion fraction and diffusion coefficient derived from different b-value combinations. *Magn Reson Med* 69:553–562
- Le Bihan D, Breton E, Lallemand D, Aubin ML, Vignaud J, Laval-Jeantet M (1988) Separation of diffusion and perfusion in intravoxel incoherent motion MR imaging. *Radiology* 168:497–505
- Le Bihan D (2008) Intravoxel incoherent motion perfusion MR imaging: a wake-up call. *Radiology* 249:748–752
- Klauss M, Lemke A, Grunberg K et al (2011) Intravoxel incoherent motion MRI for the differentiation between mass forming chronic pancreatitis and pancreatic carcinoma. *Invest Radiol* 46:57–63
- Sumi M, Van Cauteren M, Sumi T, Obara M, Ichikawa Y, Nakamura T (2012) Salivary gland tumors: use of intravoxel incoherent motion MR imaging for assessment of diffusion and perfusion for the differentiation of benign from malignant tumors. *Radiology* 263:770–777
- Sigmund EE, Vivier PH, Sui D et al (2012) Intravoxel incoherent motion and diffusion-tensor imaging in renal tissue under hydration and furosemide flow challenges. *Radiology* 263:758–769
- Shinmoto H, Tamura C, Soga S et al (2012) An intravoxel incoherent motion diffusion-weighted imaging study of prostate cancer. *AJR Am J Roentgenol* 199:W496–W500
- Riches SF, Hawtin K, Charles-Edwards EM, de Souza NM (2009) Diffusion-weighted imaging of the prostate and rectal wall: comparison of biexponential and monoexponential modelled diffusion and associated perfusion coefficients. *NMR Biomed* 22:318–325
- Wang H, Cheng L, Zhang X et al (2010) Renal cell carcinoma: diffusion-weighted MR imaging for subtype differentiation at 3.0 T. *Radiology* 257:135–143
- Rosenkrantz AB, Niver BE, Fitzgerald EF, Babb JS, Chandarana H, Melamed J (2010) Utility of the apparent diffusion coefficient for

- distinguishing clear cell renal cell carcinoma of low and high nuclear grade. *AJR Am J Roentgenol* 195:W344–W351
32. Kyriazi S, Collins DJ, Messiou C et al (2011) Metastatic ovarian and primary peritoneal cancer: assessing chemotherapy response with diffusion-weighted MR imaging—value of histogram analysis of apparent diffusion coefficients. *Radiology* 261:182–192
 33. Wang S, Kim S, Zhang Y et al (2012) Determination of grade and subtype of meningiomas by using histogram analysis of diffusion-tensor imaging metrics. *Radiology* 262:584–592
 34. Pope WB, Kim HJ, Huo J et al (2009) Recurrent glioblastoma multiforme: ADC histogram analysis predicts response to bevacizumab treatment. *Radiology* 252:182–189
 35. Peng Y, Jiang Y, Yang C et al (2013) Quantitative analysis of multiparametric prostate MR images: differentiation between prostate cancer and normal tissue and correlation with Gleason score—a computer-aided diagnosis development study. *Radiology* 267:787–796
 36. Oto A, Yang C, Kayhan A et al (2011) Diffusion-weighted and dynamic contrast-enhanced MRI of prostate cancer: correlation of quantitative MR parameters with Gleason score and tumor angiogenesis. *AJR Am J Roentgenol* 197:1382–1390
 37. DeLong ER, DeLong DM, Clarke-Pearson DL (1988) Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 44:837–845
 38. Hambrock T, Hoeks C, Hulsbergen-van de Kaa C et al (2012) Prospective assessment of prostate cancer aggressiveness using 3-T diffusion-weighted magnetic resonance imaging-guided biopsies versus a systematic 10-core transrectal ultrasound prostate biopsy cohort. *Eur Urol* 61:177–184
 39. Döpfert J, Lemke A, Weidner A, Schad LR (2011) Investigation of prostate cancer using diffusion-weighted intravoxel incoherent motion imaging. *Magn Reson Imaging* 29:1053–1058
 40. Ocak I, Bernardo M, Metzger G et al (2007) Dynamic contrast-enhanced MRI of prostate cancer at 3 T: a study of pharmacokinetic parameters. *AJR Am J Roentgenol* 189:849
 41. Zelhof B, Pickles M, Liney G et al (2009) Correlation of diffusion-weighted magnetic resonance data with cellularity in prostate cancer. *BJU Int* 103:883–888
 42. Gibbs P, Liney GP, Pickles MD, Zelhof B, Rodrigues G, Turnbull LW (2009) Correlation of ADC and T2 measurements with cell density in prostate cancer at 3.0 Tesla. *Invest Radiol* 44:572–576
 43. Woodfield CA, Tung GA, Grand DJ, Pezzullo JA, Machan JT, Renzulli JF 2nd (2010) Diffusion-weighted MRI of peripheral zone prostate cancer: comparison of tumor apparent diffusion coefficient with Gleason score and percentage of tumor on core biopsy. *AJR Am J Roentgenol* 194:W316–W322
 44. Sinha AA, Quast BJ, Reddy PK et al (2004) Microvessel density as a molecular marker for identifying high-grade prostatic intraepithelial neoplasia precursors to prostate cancer. *Exp Mol Pathol* 77:153–159
 45. Ren J, Huan Y, Wang H et al (2008) Dynamic contrast-enhanced MRI of benign prostatic hyperplasia and prostatic carcinoma: correlation with angiogenesis. *Clin Radiol* 63:153–159
 46. Schlemmer HP, Merkle J, Grobholz R et al (2004) Can pre-operative contrast-enhanced dynamic MR imaging for prostate cancer predict microvessel density in prostatectomy specimens? *Eur Radiol* 14:309–317
 47. Kiessling F, Lichy M, Grobholz R et al (2004) Simple models improve the discrimination of prostate cancers from the peripheral gland by T1-weighted dynamic MRI. *Eur Radiol* 14:1793–1801
 48. Cazares LH, Drake RR, Esquela-Kirscher A, Lance RS, Semmes OJ, Troyer DA (2010) Molecular pathology of prostate cancer. *Cancer Biomark* 9:441–459
 49. Fournier G, Valeri A, Mangin P, Cussenot O (2004) Prostate cancer. Epidemiology. Risk factors. *Pathology. Ann Urol (Paris)* 38:187–206
 50. Tozer DJ, Jager HR, Danchaivijitr N et al (2007) Apparent diffusion coefficient histograms may predict low-grade glioma subtype. *NMR Biomed* 20:49–57
 51. Kang Y, Choi SH, Kim YJ et al (2011) Gliomas: Histogram analysis of apparent diffusion coefficient maps with standard- or high-b-value diffusion-weighted MR imaging—correlation with tumor grade. *Radiology* 261:882–890
 52. Rosenkrantz AB, Sigmund EE, Johnson G et al (2012) Prostate cancer: feasibility and preliminary experience of a diffusional kurtosis model for detection and assessment of aggressiveness of peripheral zone cancer. *Radiology* 264:126–135