

OPEN

Non-Gaussian models of diffusion weighted imaging for detection and characterization of prostate cancer: a systematic review and meta-analysis

V. Brancato, C. Cavaliere*, M. Salvatore  & S. Monti 

The importance of Diffusion Weighted Imaging (DWI) in prostate cancer (PCa) diagnosis have been widely handled in literature. In the last decade, due to the mono-exponential model limitations, several studies investigated non-Gaussian DWI models and their utility in PCa diagnosis. Since their results were often inconsistent and conflicting, we performed a systematic review of studies from 2012 examining the most commonly used Non-Gaussian DWI models for PCa detection and characterization. A meta-analysis was conducted to assess the ability of each Non-Gaussian model to detect PCa lesions and distinguish between low and intermediate/high grade lesions. Weighted mean differences and 95% confidence intervals were calculated and the heterogeneity was estimated using the I^2 statistic. 29 studies were selected for the systematic review, whose results showed inconsistency and an unclear idea about the actual usefulness and the added value of the Non-Gaussian model parameters. 12 studies were considered in the meta-analyses, which showed statistical significance for several non-Gaussian parameters for PCa detection, and to a lesser extent for PCa characterization. Our findings showed that Non-Gaussian model parameters may potentially play a role in the detection and characterization of PCa but further studies are required to identify a standardized DWI acquisition protocol for PCa diagnosis.

Prostate cancer (PCa) is the second most common cancer among men¹. Accurate detection and assessment of cancerous lesion aggressiveness according to the Gleason grading system² is important for the most appropriate treatment strategy³. Currently, the most commonly used prostate cancer screening paradigm consists of the serum prostate-specific antigen (PSA) test, digital rectal examination, transrectal ultrasound (TRUS), and prostatic biopsies.

However, since these methods are often inaccurate and also invasive, there is a growing need for non-invasive tools to improve diagnosis of prostate cancer in terms of both detection and characterization.

Recently, a multiparametric Magnetic Resonance Imaging (MRI) based approach combining anatomic T1 or T2-weighted imaging with functional methods as Diffusion Weighted Imaging (DWI) and Dynamic Contrast Enhanced (DCE) imaging has significantly strengthened the role of MRI in diagnosis of PCa. There are several studies^{4–6} showing that multiparametric MRI is a useful tool that helps to detect prostate cancer foci, especially in patients with prior negative biopsy and permanently high PSA values. The classification used for evaluating the prostate as seen in MRI is known as PI-RADS (Prostate Imaging Reporting and Data System)⁷. This classification is based on a score from 1 to 5, with 1 indicating most probably benign lesions and 5 indicating lesions with a very high probability of malignancy. The use of PI-RADS, however, requires radiologists with high level of experience in MRI.

DWI technique exploits the diffusion phenomenon, which depends on the microscopic mobility of water molecules. Depending on how much water molecules movement is limited by tissue structure, the DWI signal intensity changes, and this has been proven useful in PCa to distinguish benign from malignant lesions and to characterize aggressiveness in terms of distinction between high- and low-grade tumors and correlation of tumor

IRCCS SDN, Napoli, Italy. *email: ccavaliere@sdn-napoli.it

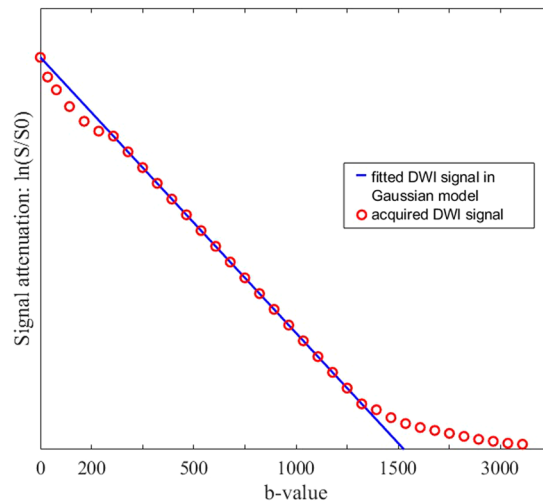


Figure 1. Example of fitted curve of diffusion weighted imaging (DWI) signal against the b -values. According to Gaussian model, when plotted against the b -values, the diffusion magnetic resonance (MR) signal ($\ln(S/S_0)$) would follow a straight line whose slope is the ADC (apparent diffusion coefficient). Non-Gaussian DWI models were introduced to describe the deviation of measured data from this expected line. Abbreviations: DWI, diffusion weighted imaging.

with Gleason Score (GS)^{8–14}. Not only correlation with GS is important but it is of clinical importance to separate low Gleason grade PCa lesions from intermediate and high Gleason grade lesions¹⁵.

Several diffusion models have been studied in the field of PCa. Gaussian model has been largely used for cancer detection and characterization and has allowed meaningful results to be achieved. ADC values of PCa are generally lower than those of prostatic Normal Tissue (NT)^{8,16} and also reflect tumor aggressiveness, showing a negative correlation with GS^{9,17}. However, the assumption at the basis of conventional DWI are not always accurate: water molecules experience vastly different environments in tissues, so *in vivo* water diffusion is much more complicated and often presents Non-Gaussian behavior. In terms of signal intensity, at low b -values (≤ 200 s/mm²) the signal attenuation is greater than expected, while it is lower at larger b -values (≥ 1500 s/mm²) (Fig. 1). Consequently, Non-Gaussian diffusion models have been proposed to better describe diffusion signal behavior, which can be directly related to tissue physiologic and pathologic characteristic. The most commonly used models are the Intravoxel Incoherent Motion (IVIM), the Diffusion Kurtosis Imaging (DKI), the Biexponential (BE), and the Stretched Exponential (SE).

There is a significant amount of studies aiming to investigate diagnostic performances of Non-Gaussian models, but their results suffer from inconsistency, insignificance and a not clear physical interpretation of Non-Gaussian parameters.

On this basis, the aim of our review is to evaluate the diagnostic performance of Non-Gaussian DWI models in terms of detection and characterization of prostate cancer (PCa). Specifically, after a detailed assessment, selection and systematic review of studies examining both Non-Gaussian and standard Gaussian DWI models, which will be shortly introduced in the following subsection, published data related to the most commonly used models parameters were statistically analyzed to evaluate their ability to detect PCa lesions and distinguish between low grade lesions ($GS \leq 6$) and intermediate/high grade lesions ($GS \geq 7$).

Diffusion Weighted MRI Mathematical Models in PCa

The most commonly used mathematical models used to fit DWI signal are listed below:

1. Mono-exponential Model (ME)¹⁸

$$\frac{S(b)}{S_0} = \exp(-b \cdot ADC)$$

2. Intravoxel Incoherent Motion Model (IVIM)¹⁹

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

3. Biexponential Model (BE)^{20–22}

$$\frac{S(b)}{S_0} = f_{slow} \cdot \exp(-b \cdot D_{slow}) + f_{fast} \cdot \exp(-b \cdot D_{fast})$$

4. Stretched Exponential model (SE)²³

Classification based on Non-Gaussian models	Classification based on PCa diagnosis
Intravoxel Incoherent Motion Model (IVIM)	PCa detection
Biexponential Model (BE)	Characterization of PCa aggressiveness
Stretched Exponential Model (SE)	
Diffusion Kurtosis Imaging (DKI)	

Table 1. Article classifications according to the Non-Gaussian model examined and the diagnostic purpose of the study. Abbreviations: PCa, prostate cancer.

$$\frac{S(b)}{S_0} = \exp[-(b \cdot DDC)^\alpha]$$

5. Diffusion Kurtosis Imaging (DKI)²⁴

$$\frac{S(b)}{S_0} = \exp\left(-b \cdot D_K + \frac{b^2 \cdot D_K^2 \cdot K}{6}\right)$$

In all the equations, $S(b)$ is the measured signal intensity at a certain b , S_0 the signal intensity at $b=0$ and b is a factor that measures the degree of diffusion weighting applied. In the ME model, *i.e.* the standard gaussian model, ADC is the Apparent Diffusion Coefficient, an average value related to diffusion. In the IVIM model f is the perfusion fraction, D the molecular diffusion coefficient and D^* the pseudo-diffusion coefficient. D_{fast} and D_{slow} are, respectively, fast and slow diffusion coefficients of the BE model and f_{fast} and f_{slow} their amplitudes. DDC is the distributed diffusion coefficient of the SE model, and α is the heterogeneity index. Finally, in the DKI model, D_K is the diffusion coefficient corrected for kurtosis, and K is the kurtosis coefficient. See Supplementary Material-S1- for more details on Non-Gaussian DWI models.

Methods

Search strategy and selection criteria. A systematic search for relevant published studies examining and, if any, comparing mono-exponential DWI and Non-Gaussian DWI models in PCa diagnosis was conducted. The most relevant scientific electronic databases (PubMed, Cochrane Library, MEDLINE, ScienceDirect, Google Scholar) were comprehensively explored and used to build the search. Only studies published after 2012 were selected. The search strategy included the key terms listed in Supplementary materials -S2-. The literature search was restricted to English language publications and studies of human subject.

Two reviewers, after having independently screened identified titles and abstracts, assessed the full text of the articles that evaluated mono-exponential DWI and at least one Non-Gaussian model, between IVIM, BE, SE, DKI, for diagnosis of PCa (detection and/or characterization of aggressiveness) and were not review articles. For articles meeting these criteria with full text available, the following further selection criteria had to be fulfilled: presence of PCa histopathological confirmation (either from biopsy or radical prostatectomy), of information about DW-MRI protocol, of a maximum b -value at least equal to 1500 s/mm^2 (if DKI was performed). Moreover, articles were excluded if computed b -values were used, if they concerned only model fitting quality or differentiation between PCa and benign lesions, not evaluating NT. The above-mentioned selection procedure was employed to perform a systematic review, providing a qualitative synthesis of currently used Non-Gaussian DWI models in PCa diagnosis.

In order to perform a meta-analysis, a further exclusion procedure was performed: studies were removed if they focused only on correlation of parameters with GS, and not on their ability to differentiate low-GS from high-GS tumors; if number of analyzed region of interest (ROI) was not mentioned; if data were reported in the form of mean and Inter Quartile Range (IQR), median and IQR, or median and 95% Confidence Interval (CI); if mean and standard deviation of parameters were not reported (or could not be calculated); if $b=0 \text{ mm}^2/\text{s}$ was not included in the DWI protocol. Moreover, if there was a high heterogeneity in fitting functions/procedures used for biexponential models, statistical analyses had not been conducted.

Planning of the study. The articles were classified according to the Non-Gaussian models examined and the diagnostic purpose they had, as reported in Table 1.

Consequently, our work was organized in accordance with Fig. 2: for each Non-Gaussian model, a qualitative analysis followed by a meta-analysis (when a sufficient number of articles was available after the application of selection criteria) was performed both to assess the difference of the mean value of IVIM, BE, DKI and SE parameters between NT and PCa (PCa detection) and to assess the parameters mean differences between low GS and intermediate/high GS PCa (characterization of PCa aggressiveness).

Meta-analyses methods. Meta-analyses were conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement²⁵ (See Supplementary Materials-S3-for PRISMA Checklist).

Although there were not enough data for performing an assessment of diagnostic accuracy, the quality of studies included in meta-analysis was evaluated, using the QUADAS-2²⁶ tool included in RevMan (version 5.3,

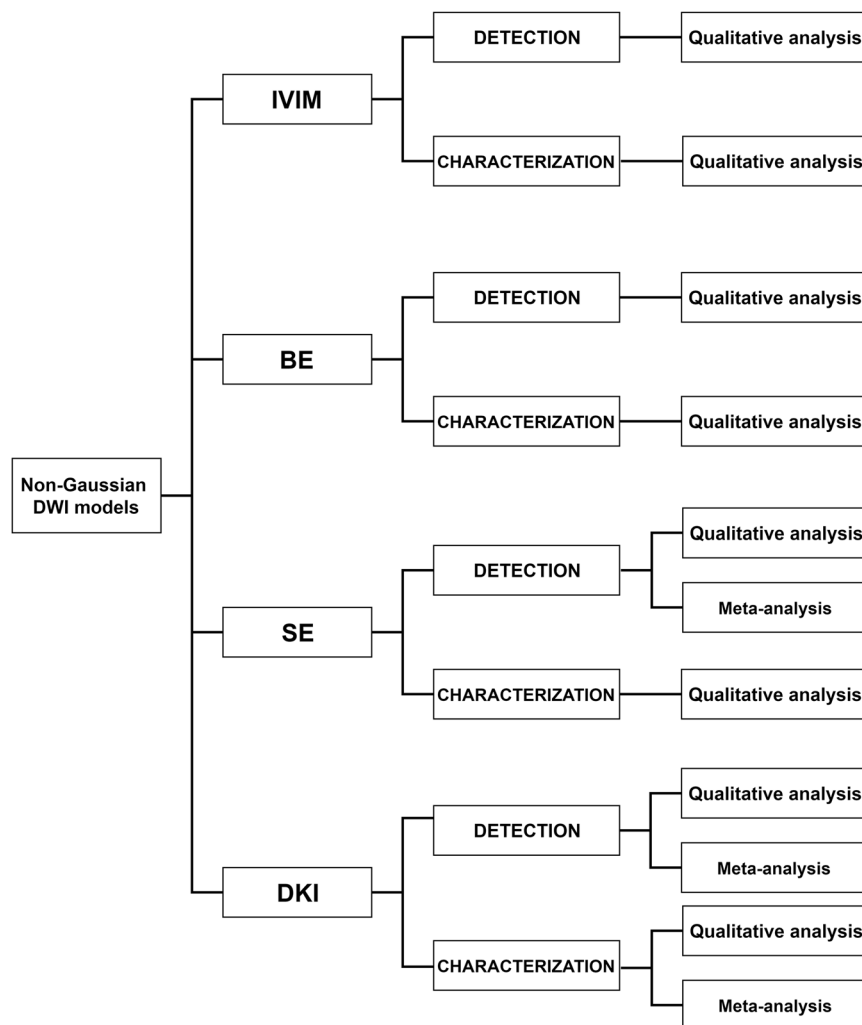


Figure 2. Scheme reporting planning of the study. Abbreviations: DWI, Diffusion Weighted Imaging; IVIM, Intravoxel Incoherent Motion model; BE, Biexponential model; SE, Stretched Exponential model; DKI, Diffusion Kurtosis Imaging.

The Cochrane Collaboration). The quality of each study was evaluated by two reviewers independently and any disagreement was resolved by consensus. If meta-analysis included a number of studies superior to 10^{27,28}, publication bias was assessed by visually inspecting a funnel plot.

Mean and Standard Deviation (SD) of diffusion model parameters were extracted from each selected article. To analyze the differences between groups, Weighted Mean Differences (WMD) and 95% CI were calculated. The overall result was considered statistically significant if the test for overall effect returns a probability value lower than 0.05.

Heterogeneity among included studies was assessed using the Q statistic of the chi-square value test and the inconsistency index of Higgins²⁹ I², with values of 25%, 50%, and 75% considered as low, moderate, and high, respectively. If the P-value of heterogeneity test was less than 0.1 or the I²-value was greater than 50%, the summary estimate was analyzed by a random-effects model³⁰. Otherwise, a fixed-effects model was applied.

All statistical computations were performed using RevMan (version 5.3, The Cochrane Collaboration).

Results

The PRISMA flow diagram of included studies according to the inclusion and exclusion criteria is reported in Fig. 3.

Results of qualitative analysis. 29 studies fulfill all the inclusion criteria and were involved in the qualitative analysis: their characteristics are summarized in Tables 2 and 3.

In all selected studies, ADC value proved to be a useful tool for discriminating both NT from PCa and low- from high-GS tumors, with lower values in tumors than in NT and in high-GS tumors than in low-GS tumors. On the other hand, findings on Non-Gaussian model features are not always equally consistent.

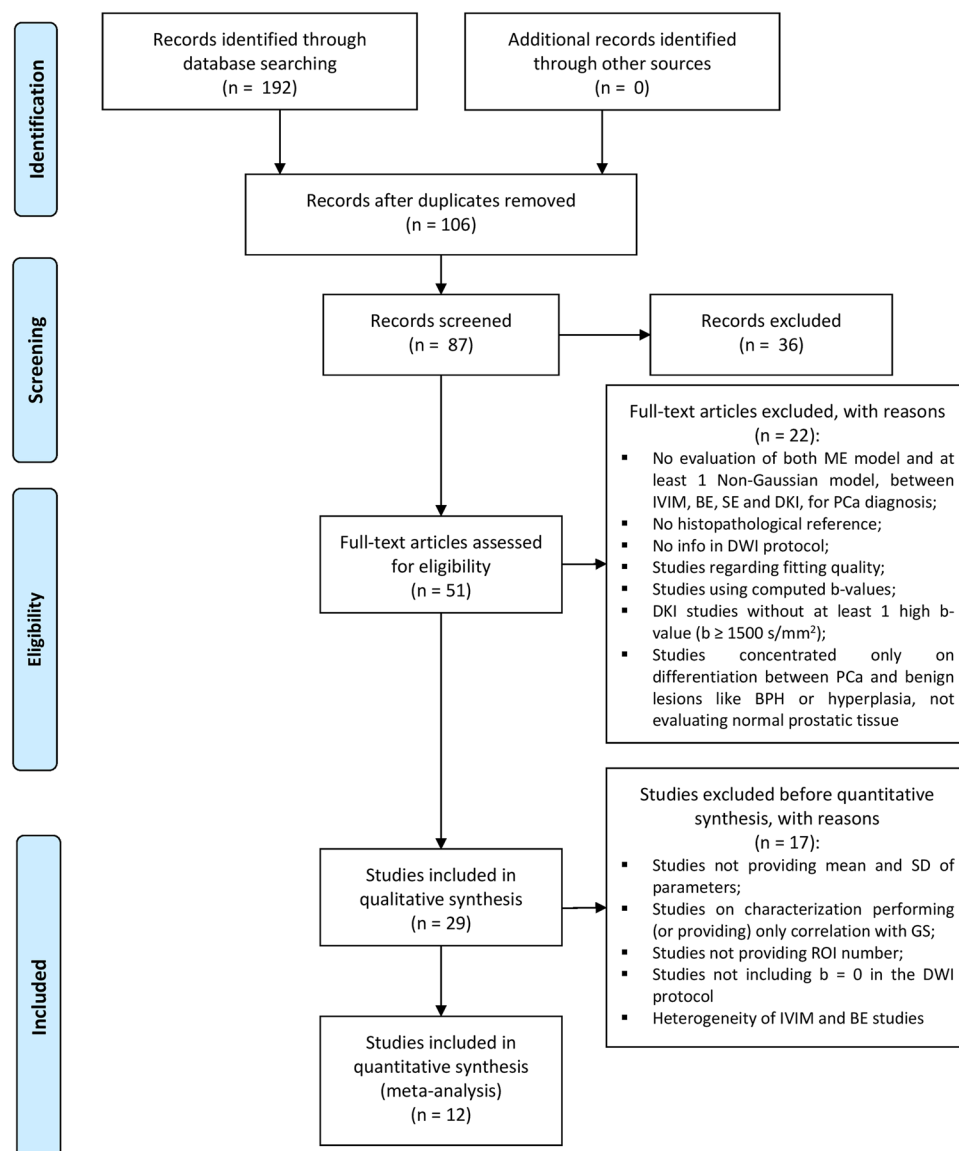


Figure 3. PRISMA flow diagram of the study selection procedure. Abbreviations: DWI, Diffusion Weighted Imaging; IVIM, Intravoxel Incoherent Motion model; BE, Biexponential model; SE, Stretched Exponential model; DKI, Diffusion Kurtosis Imaging; BPH, benign prostatic hypertrophy; SD, standard deviation; GS, Gleason Score; ROI, region of interest.

Studies on IVIM. 14 studies included IVIM comparison with ADC: 6 on detection^{31–36}, 4 on characterization^{37–40}, and 4 on both^{41–44}. Most studies on detection found that D was significantly smaller in PCa when compared to NTs^{31–36,41–44}, but Mazzoni *et al.*³⁴ revealed a dependence on b-value range, showing statistical significance only using a maximum b value of 800 s/mm^2 . Results on f, when significant, revealed a lower value in PCa than in NT^{31–34,36,43}, with the exception of study by Ueda *et al.*³⁵ who, conversely, found f significantly higher in peripheral zone (PZ) cancer than in normal PZ tissue, although they found no significance when performing the same analysis in transitional zone (TZ). Results on D* were found to be not significant in many studies^{33–36}. Pesapane *et al.*⁴³ found that D* was significantly higher in tumor tissue when compared to NT, in accordance with Kuru *et al.*⁴¹ and Valerio *et al.*⁴². While Linear Discriminant Analysis performed by Pesapane *et al.*⁴³ and Valerio *et al.*⁴² showed that the additional use of IVIM increased performances of conventional T2/DWI for PCa detection, ROC analysis performed by Kuru *et al.* proved that none of IVIM parameters yielded a clear added value, such as in Feng *et al.*³⁶. Whereas, Martin *et al.*³² found D and f to be statistically significant, obtaining also a slightly higher AUC for these two parameters than for ADC. In PCa aggressiveness characterization, D seems to be the most performing IVIM parameter^{37–42}, resulting lower in high- than in low-GS tumors in accordance with ROC analysis performed by Yang *et al.*³⁸. About D*, only Valerio *et al.*⁴² found its value to be significantly higher in high- than in low-GS PCa, while f was considered unable to discriminate between GS grade in all selected studies. Pesapane *et al.*⁴³ identified all three IVIM parameters as not useful for characterization on PCa aggressiveness. In addition, some studies carried out a correlation analysis between the IVIM parameters and the GS: they revealed

Author	Year	Non-Gaussian Models	Purpose (D/C)	Sd	No. of PCa patients	No. of PCa regions	No. of NT regions	No. of Low GS regions (<=6)	No. of High GS regions (>6)	Included in MA	Reasons for exclusion from MA
Shinmoto <i>et al.</i> ⁴¹	2012	IVIM	D	R	26	N/E	N/E	—	—	No	Study on IVIM model
Liu <i>et al.</i> ⁴⁵	2013	BE	D	P	23	N/E	N/E	—	—	No	Study on BE model
Kuru <i>et al.</i> ⁴¹	2014	IVIM	D, C	R	27	N/E	N/E	N/E	N/E	No	Study on IVIM model
Zhang <i>et al.</i> ³⁷	2015	IVIM	C	R	48	—	—	N/E	N/E	No	Study on IVIM model
Martin <i>et al.</i> ³²	2014	IVIM	D	R	36	N/E	N/E	—	—	No	Study on IVIM model
Valerio <i>et al.</i> ⁴²	2016	IVIM	D, C	P	53	N/E	N/E	N/E	N/E	No	Study on IVIM model
Yang <i>et al.</i> ³⁸	2016	IVIM	C	R	41	—	—	N/E	N/E	No	Study on IVIM model
Barbieri <i>et al.</i> ³⁹	2017	IVIM	C	P	84	—	—	N/E	N/E	No	Only analysis of correlation with GS reported
Bao <i>et al.</i> ⁴⁰	2017	IVIM	C	P	30	—	—	N/E	N/E	No	Only analysis of correlation with GS reported
Pesapane <i>et al.</i> ⁴³	2017	IVIM	D, C	P	31	N/E	N/E	N/E	N/E	No	Study on IVIM model
Liu <i>et al.</i> ⁴⁹	2015	SE	D	P	27	31	62	—	—	Yes	
Liu <i>et al.</i> ⁵⁰	2018	SE	C	R	75	—	—	N/E	N/E	Yes	
Rosenkrantz <i>et al.</i> ⁵⁴	2012	DKI	D, C	R	47	121	47	51	70	Yes	
Tamura <i>et al.</i> ⁵¹	2014	DKI	D	R	20	24	20	—	—	Yes	
Suo <i>et al.</i> ⁵⁵	2014	DKI	D, C	R	19	19	19	9	10	Yes	
Roethke <i>et al.</i> ⁵⁶	2015	DKI	D, C	R	55	55	55	12	43	Yes	
Wang <i>et al.</i> ⁵²	2015	DKI	C	R	110	—	—	49	77	Yes	
Tamada <i>et al.</i> ⁵⁷	2017	DKI	D, C	R	285	285	285	73	311	Yes	
Wang <i>et al.</i> ⁵³	2018	DKI	C	R	67	—	—	N/E	N/E	No	b = 0 not included in DWI protocol
Quentin <i>et al.</i> ³³	2012	IVIM	D	P	8	N/E	N/E	—	—	No	Study on IVIM model
Mazzoni <i>et al.</i> ³⁴	2014	IVIM, DKI	D	R	57	73	45	—	—	Yes	Exclusion of IVIM model
Ueda <i>et al.</i> ³⁵	2016	IVIM	D	R	63	64	64	—	—	No	Study on IVIM model
Jambor <i>et al.</i> ⁴⁶	2015	BE, DKI, SE	D	P	16	N/E	N/E	—	—	No	Mean - SD values not reported
Toivonen <i>et al.</i> ⁴⁸	2015	BE, DKI, SE	D, C	P	50	N/E	N/E	N/E	N/E	No	Mean - SD values not reported
Feng <i>et al.</i> ³⁶	2017	IVIM, DKI, SE	D	P	56	138	198	—	—	Yes	Exclusion of IVIM model
Barrett <i>et al.</i> ⁵⁸	2017	DKI	D, C	P	30	N/E	N/E	N/E	N/E	No	Mean - SD values not reported
Merisaari <i>et al.</i> ⁴⁴	2017	IVIM, SE	D, C	P	81	N/E	N/E	N/E	N/E	No	Mean - SD values not reported
Mazaheri <i>et al.</i> ⁴⁷	2018	BE, DKI, SE	D	R	55	55	55	—	—	Yes	Exclusion of BE model
Langkilde <i>et al.</i> ¹⁵	2018	BE, DKI, SE	D, C	R	40	40	111	N/A	N/A	Yes	Exclusion of BE model

Table 2. Selected studies for qualitative and quantitative analysis. Abbreviations: IVIM, Intravoxel Incoherent Motion Model; BE, Biexponential Model; DKI, Diffusion Kurtosis Imaging; SE, Stretched Exponential Model; D, detection; C, characterization; Sd, study design; P, prospective; R, retrospective; PCa, prostate cancer; NT, normal tissue; GS, Gleason Score; MA, meta-analysis; SD, standard deviation; N/E, not extracted because not necessary for quantitative analysis; —, non-existent. Notes: the number of PCa regions, Low GS regions and High GS regions were reported as sum of all PCa, Low GS or High GS regions respectively, regardless of the affected prostate zone (peripheral zone, transition zone, central gland); the number of NT regions were reported as sum of all NT regions, regardless of the affected prostate zone and the patient on which the ROI was placed (e.g. healthy volunteer, PCa patient). The column “Reason for exclusion from MA” reported the reason why the study under investigation, or which analyzed model, was excluded from the meta-analysis. As regards IVIM and BE models, they were excluded from meta-analysis due to heterogeneity in fitting functions/procedures among selected studies and this was indicated by the sentences “Study on IVIM model” or “Study on BE model”, if the study included only these models, “Exclusion of IVIM model” or “Exclusion of BE model”, if the study reported other Non-Gaussian DWI models that were, instead, retained into meta-analysis.

a significant negative correlation between D and GS^{38–40}, while no correlation was found for f and D*. Merisaari *et al.*⁴⁴ used machine learning techniques in order to evaluate if GS prediction could be improved using ADC in combination with IVIM, but they concluded that ADC continued to be the best-performing parameter for this scope.

Studies on BE. 5 studies included BE comparison with ADC: 4 on detection^{15,45–47} and 1 on both detection and characterization⁴⁸. Liu *et al.*⁴⁵ found all three BE parameters significantly lower in PCa than in NT, in accordance with Mazaheri *et al.* and Langkilde *et al.*^{15,47}, who found similar results for D_{slow} and f_{fast} but any significance for D_{fast} . In contrast, Jambor *et al.*⁴⁶ although identified in BE the best performance for normal and PCa data fitting, found median values of its three parameters not reaching statistical significance for PCa detection. Toivonen *et al.*⁴⁸ evaluated BE model in terms of both detection and characterization, performing two ROC analysis and a correlation analysis with GS. Their results did not demonstrate improvements in PCa detection and characterization compared with mono-exponential model and only f showed a significant negative correlation with GS.

Studies on SE. 8 studies included SE comparison with ADC: 5 on detection^{15,36,46,47,49}, 2 on characterization^{44,50}, and 1 on both⁴⁸. According to Mazaheri *et al.*⁴⁷ and Liu *et al.*⁴⁹, DDC and α in PCa were significantly lower than in NT, showing that SE model parameters could provide additional information to ME model. These results are consistent with those of Feng *et al.*³⁶ who, however, also performed a ROC analysis showing good performances for both parameters, but no significant performances improvement with respect to ADC. Jambor *et al.*⁴⁶ found similar results for DDC, but not for α . Even ROC analysis performed by Toivonen *et al.*⁴⁸ reveals significance for DDC, both in PCa detection and characterization, showing also a negative correlation between DDC and GS, even if performances are not such to substitute the ME model. In the study by Liu *et al.*⁵⁰ DDC was found to be the only SE parameter able to differentiate low- from high-grade tumors, showing a significant negative correlation with GS. Merisaari *et al.*⁴⁴ found a statistically significant correlation for both DDC and α with GS (negative and positive respectively), a high AUC for low- versus high-GS classification and an AUC improvement combining DDC and α by means of machine learning. Nevertheless, none of these parameters or their combination considerably outperformed ADC parameter of ME model.

Studies on DKI. 15 studies included DKI comparison with ADC. Of these, 6 were on detection^{34,36,46,47,51}, 2 on characterization^{52,53}, and 7 dealt with both together^{15,48,54–58}.

All the studies revealed a significantly higher K value in tumor tissue than in NT, with increasing trend in high GS lesions, but there were different opinions regarding the added value offered by the use of DKI for PCa diagnosis. Preliminary findings by Rosenkrantz *et al.*⁵⁴ indicated higher capability of DKI model for both distinguishing benign from malignant PCa lesions and low- from high-grade PCa lesions when compared to ADC, by means of ROC analyses. They found K significantly higher and D_K significantly lower in PCa than in NT, and the same behavior was observed when low- and high- GS lesions were compared. The findings on PCa detection were successively confirmed by Tamura *et al.*⁵¹, Jambor *et al.*⁴⁶, Suo *et al.*⁵⁵, Mazaheri *et al.*⁴⁷, and Mazzoni *et al.*³⁴ who found that K was the best-performing parameter. Feng *et al.*³⁶, Roethke *et al.*⁵⁶, and Langkilde *et al.*¹⁵, although their statistical significance test results were in accordance with the above-mentioned studies, revealed not sufficiently higher performance of D_K or K compared to ADC at ROC analysis, in accordance with Toivonen *et al.*⁴⁸. A similar conclusion was made by Tamada *et al.*⁵⁷, who studied only K parameter. In the last 4 papers the authors, together with Wang *et al.*⁵³, found similar results also in PCa characterization. Suo *et al.*⁵⁵ found significant differences among low- and high-grade lesions for K value, but not for D_K value. Conversely, results by Barrett *et al.*⁵⁸ were not significant, showing poor capability for both D_K and K at separating low- and high-grade lesions. Correlation analysis for the DKI parameters, evaluating their relationship with GS, revealed a significantly positive correlation between K and GS^{48,52,53,55}, and a significantly negative correlation between D_K and GS^{48,52}. Suo *et al.*⁵⁵, however, found no significant correlation between D_K value and GS.

Results of meta-analyses. In order to perform the meta-analyses indicated in Fig. 2, 17 studies were fully excluded from previous selection. Data from the same study were considered multiple times in the following cases: different b-value ranges were used to perform acquisition/processing of DW-MRI images; PCa tissue was compared with different zones of NT (e.g. PCa vs normal TZ and PCa vs normal PZ); NT was compared with different PCa zones (e.g. NT vs PCa in TZ and NT vs PCa in PZ); NT was compared with PCa with different GS (e.g. NT vs PCa with GS = 6 and NT vs PCa with GS ≥ 7); low grade PCa was compared to more than one PCa with higher GS. Data from the study of Langkilde *et al.*¹⁵ were extracted only for the analyses including $b = 0 \text{ s/mm}^2$.

Finally, as for PCa detection, studies concerning DKI and SE were respectively 32, and 10. For the SE, 1770 regions (694 positive) were analyzed; for the DKI, 5167 regions (2136 positive) were analyzed for K and 3331 (1439 positive) for D_K .

As for characterization, the meta-analysis was performed only for DKI model, because there were no sufficient SE data available. DKI was addressed by 9 studies, with 897 lesions (300 low-grade) included for K and 440 (154 low-grade) for D_K .

The overall quality of included studies in meta-analyses, for both PCa detection and characterization, was considered good for our purposes (See Supplementary Materials –S4–).

Concerning IVIM and BE models, the high heterogeneity in applied fitting functions and procedures (see Table 2) led us to decide not to perform any quantitative analysis on these models.

Performance of models in PCa detection. The random-effect analysis was used for all SE and DKI parameters. The heterogeneity was high ($I^2 > 50\%$ and $P < 0.0001$). D_K was significantly higher in NT, while K was significantly higher in PCa tissue. These results are highlighted by the forest plot in Figs 4 and 5.

Performance of models in PCa characterization. D_K was found to be significantly higher in PCa lesions with low GS, while K was significantly higher in those with high GS. These results are highlighted by the forest plot in Fig. 6.

Author	FS [T]	TR/TE [ms/ms]	Seq.	b-values [s/mm ²]	Non-Gaussian fitting function(s)	Non-Gaussian Parameters	Fitting procedure	Initialization values	Methods to prevent finding local minima
Shinmoto <i>et al.</i> ³¹	3.0	5132/40	NR	0, 10, 20, 30, 50, 80, 100, 200, 400, 1000	$S(b)/S_0 = (1-f) \cdot \exp(-b \cdot D) + f \cdot \exp(-b \cdot (D^* + D))$	D, D*, f	NR	NR	NR
Liu <i>et al.</i> ⁴⁵	3.0	4000/71.9	SS-EPI	For ADC: 0, 1000; For BE: 0, 300, 600, 900, 1200, 1500, 1800, 2100, 2400, 2700, 3000	$S(b)/S_0 = f_{\text{slow}} \cdot \exp(-b \cdot D_{\text{slow}}) + f_{\text{fast}} \cdot \exp(-b \cdot D_{\text{fast}})$	$D_{\text{slow}}, D_{\text{fast}}, f$	NR	NR	NR
Kuru <i>et al.</i> ⁴¹	3.0	3100/52	SS-EPI	0, 50, 100, 150, 200, 250, 800	$S(b)/S_0 = (1-f) \cdot \exp(-b \cdot D) + f \cdot \exp(-b \cdot (D^* + D))$	M1: D, D*, f M2: D, f	M1: SLF for D and f, NLRf for D*; M2: BeF using LMa with D* fixed to 20 $\mu\text{m}^2/\text{ms}$	NR	NR
Zhang <i>et al.</i> ³⁷	3.0	6000/72	SS-EPI	0, 50, 150, 300, 600, 900	$S(b)/S_0 = (1-f) \cdot \exp(-b \cdot D) + f \cdot \exp(-b \cdot D^*)$	D, D*, f	BeF using LMa	NR	NR
Martin <i>et al.</i> ³²	3.0	5000/54	SS-EPI	0, 20, 40, 100, 300, 500, 1000, 2000	$S(b)/S_0 = (1-f) \cdot \exp(-b \cdot D) + f \cdot \exp(-b \cdot (D^* + D))$	D, f	NR	NR	NR
Valerio <i>et al.</i> ⁴²	3.0	3100/102	NR	For ADC: 0, 500, 1000, 3000; For IVIM: 0, 10, 20, 30, 40, 50, 80, 100, 200, 400, 800	$S(b)/S_0 = (1-f) \cdot \exp(-b \cdot D) + f \cdot \exp(-b \cdot D^*)$	D, D*, f	NR	D: $[0-10] \cdot 10^{-3} \text{ mm}^2/\text{s}$ D*: $[10-150] \cdot 10^{-3} \text{ mm}^2/\text{s}$ f: $[0-1]$	NR
Yang <i>et al.</i> ³⁸	3.0	5000/90	SS-EPI	0, 10, 20, 50, 100, 200, 500, 800	$S(b)/S_0 = (1-f) \cdot \exp(-b \cdot D) + f \cdot \exp(-b \cdot D^*)$	D, D*, f	BeF using LMa	NR	NR
Barbieri <i>et al.</i> ³⁹	3.0	2600/58	SS-EPI	0, 10, 20, 50, 130, 270, 500, 900	$S(b)/S_0 = (1-f) \cdot \exp(-b \cdot D) + f \cdot \exp(-b \cdot D^*)$	D, D*, f	Bpb	NR	NR
Bao <i>et al.</i> ⁴⁰	3.0	6800/98	SS-EPI	0, 50, 100, 150, 200, 500, 1000	$S(b)/S_0 = (1-f) \cdot \exp(-b \cdot D) + f \cdot \exp(-b \cdot D^*)$	D, D*, f	NR	NR	NR
Pesapane <i>et al.</i> ⁴³	1.5	7000/10	SS-EPI	For ADC: 0, 1000, 2000; For IVIM: 0, 10, 20, 30, 50, 80, 100, 200, 400, 800	$S(b)/S_0 = (1-f) \cdot \exp(-b \cdot D) + f \cdot \exp(-b \cdot (D^* + D))$	D, D*, f	NR	NR	NR
Liu <i>et al.</i> ⁴⁹	3.0	4000/71.9	SS-EPI	0, 500, 1000, 2000	$S(b)/S_0 = \exp[-(b \cdot \text{DDC})^\alpha]$	DDC, α	NR	—	NR
Liu <i>et al.</i> ⁵⁰	3.0	4000/71.9	SS-EPI	0, 500, 1000, 2000	$S(b)/S_0 = \exp[-(b \cdot \text{DDC})^\alpha]$	DDC, α	NR	—	NR
Rosenkrantz <i>et al.</i> ⁵⁴	3.0	3500/81	SS-EPI	0, 500, 1000, 1500, 2000	$S(b)/S_0 = \exp(-b \cdot D_K + b^2 \cdot D_K^2 \cdot K/6)$	D_K, K	NR	—	NR
Tamura <i>et al.</i> ⁵¹	3.0	5000/49	SS-EPI	0, 10, 20, 30, 50, 80, 100, 200, 400, 1000, 1500	$S(b)/S_0 = \exp(-b \cdot D_K + b^2 \cdot D_K^2 \cdot K/6)$	D_K, K	NR	—	NR
Suo <i>et al.</i> ⁵⁵	3.0	3940/106	SS-EPI	0, 500, 800, 1200, 1500, 2000	$S(b)/S_0 = \exp(-b \cdot D_K + b^2 \cdot D_K^2 \cdot K/6)$	D_K, K	NLLS	—	NR
Roethke <i>et al.</i> ⁵⁶	3.0	For ADC: 3100/52 For DKI: 2700/70	SS-EPI	For ADC: 0, 800 For DKI: 0, 50, 250, 500, 750, 1000, 1250, 1500, 2000	$S(b)/S_0 = \exp(-b \cdot D_K + b^2 \cdot D_K^2 \cdot K/6)$	D_K, K	LMa	—	NR
Wang <i>et al.</i> ⁵²	3.0	6800/98	SS-EPI	0, 700, 1400, 2100	$S(b)/S_0 = \exp(-b \cdot D_K + b^2 \cdot D_K^2 \cdot K/6)$	D_K, K	NR	—	NR
Tamada <i>et al.</i> ⁵⁷	3.0	3500/81	SS-EPI	For ADC: 0, 1000 For DKI: 0, 500, 1000, 1500, 2000	$S(b)/S_0 = \exp(-b \cdot D_K + b^2 \cdot D_K^2 \cdot K/6)$	K	NR	—	NR
Wang <i>et al.</i> ⁵³	3.0	4500/95	SS-EPI	200, 500, 1000, 1500, 2000	$S(b)/S_0 = \exp(-b \cdot D_K + b^2 \cdot D_K^2 \cdot K/6)$	K	NR	—	NR
Quentin <i>et al.</i> ³³	3.0	2600/89	SS-EPI	0, 50, 100, 150, 200, 300, 400, 500, 600, 700, 800	$S(b)/S_0 = (1-f) \cdot \exp(-b \cdot D) + f \cdot \exp(-b \cdot (D^* + D))$	D, D*, f	NR	NR	NR
Mazzoni <i>et al.</i> ³⁴	3.0	2100/69	SS-EPI	0, 50, 100, 150, 200, 250, 400, 650, 800, 1000, 1400, 1800, 2300 Different ranges used: 0–2300 (group A); 0–1800 (group B); 0–800 (group C)	IVIM: $S(b)/S_0 = (1-f) \cdot \exp(-b \cdot D) + f \cdot \exp(-b \cdot D^*)$ DKI: $S(b)/S_0 = \exp(-b \cdot D_K + b^2 \cdot D_K^2 \cdot K/6)$	D, D*, f, D_K, K	NR	NR	NR
Continued									

Author	FS [T]	TR/TE [ms/ms]	Seq.	b-values [s/mm ²]	Non-Gaussian fitting function(s)	Non-Gaussian Parameters	Fitting procedure	Initialization values	Methods to prevent finding local minima
Ueda <i>et al.</i> ³⁵	3.0	4000/65	SS-EPI	0, 50, 100, 200, 500, 1000, 2000, 3000	$S(b)/S_0 = (1-f) \cdot \exp(-b \cdot D) + f \cdot \exp(-b \cdot D^*)$	D, D^*, f	SLF for D	NR	NR
Jambor <i>et al.</i> ⁴⁶	3.0	3141/51	SS-EPI	For HV: 0, 50, 100, 200, 350, 500, 650, 800, 950, 1100, 1250, 1400, 1550, 1700, 1850, 2000 For PCa patients: 0, 100, 300, 500, 700, 900, 1100, 1300, 1500, 1700, 1900, 2000	BE: $S(b)/S_0 = f_{\text{slow}} \cdot \exp(-b \cdot D_{\text{slow}}) + f_{\text{fast}} \cdot \exp(-b \cdot D_{\text{fast}})$ DKI: $S(b)/S_0 = \exp(-b \cdot D_K + b^2 \cdot D_K^2 \cdot K/6)$	$D_{\text{slow}}, D_{\text{fast}}, f_{\text{slow}}, f_{\text{fast}}, D_K, K, \text{DDC}, \alpha$	NR	For PCa: $D_{\text{fast}}: 1.0-9.0$ (ss 0.2) $\mu\text{m}^2/\text{ms}$ $D_{\text{slow}}: 0.0-4.0$ (ss 0.02) $\mu\text{m}^2/\text{ms}$ $f: 0.2-1.0$ (ss 0.1) For HV: $D_{\text{fast}}: 1.0-7.0$ (ss 0.1) $\mu\text{m}^2/\text{ms}$ $D_{\text{slow}}: 0.0-2.0$ (ss 0.01) $\mu\text{m}^2/\text{ms}$ $f: 0.2-1.0$ (ss 0.1)	Multiple initialization values
Toivonen <i>et al.</i> ⁴⁸	3.0	3141/51	SS-EPI	0, 100, 300, 500, 700, 900, 1100, 1300, 1500, 1700, 1900, 2000	BE: $S(b)/S_0 = f_{\text{slow}} \cdot \exp(-b \cdot D_{\text{slow}}) + f_{\text{fast}} \cdot \exp(-b \cdot D_{\text{fast}})$ DKI: $S(b)/S_0 = \exp(-b \cdot D_K + b^2 \cdot D_K^2 \cdot K/6)$	$D_{\text{slow}}, D_{\text{fast}}, f_{\text{slow}}, f_{\text{fast}}, D_K, K, \text{DDC}, \alpha$	BeF using LMa	$D_{\text{fast}}: 1.0-9.0$ (ss 0.2) $\mu\text{m}^2/\text{ms}$ $D_{\text{slow}}: 0.0-4.0$ (ss 0.02) $\mu\text{m}^2/\text{ms}$ $f: 0.2-1.0$ (ss 0.1)	Multiple initialization values
Feng <i>et al.</i> ³⁶	3.0	2500/84.1	SS-EPI	0, 20, 50, 80, 100, 150, 200, 400, 600, 800, 1000, 1200, 1500, 1800, 2000, 2400, 2800, 3200, 3600, 4000, 4500 Different ranges used: 0–1000; 0–2000; 0–3200; 0–4500	DKI: $S(b)/S_0 = \exp(-b \cdot D_K + b^2 \cdot D_K^2 \cdot K/6)$ SE: $S(b)/S_0 = \exp[-(b \cdot \text{DDC})^\alpha]$	$D, D^*, f, D_K, K, \text{DDC}, \alpha$	BeF using LMa	NR	NR
Barrett <i>et al.</i> ⁵⁸	3.0	For DWI: 4000/70–75 For DKI: 6000/94	SS-EPI	For ADC: 0, 150, 1000, 1400; For DKI: 0, 150, 450, 800, 1150, 1500	$S(b)/S_0 = \exp(-b \cdot D_K + b^2 \cdot D_K^2 \cdot K/6)$	D_K, K	NR	—	NR
Merisaari <i>et al.</i> ⁴⁴	3.0	1394/44	SS-EPI	0, 2, 4, 6, 9, 12, 14, 18, 23, 28, 50, 100, 300, 500	IVIM: $S(b)/S_0 = (1-f) \cdot \exp(-b \cdot D) + f \cdot \exp(-b \cdot D^*)$ SE: $S(b)/S_0 = \exp[-(b \cdot \text{DDC})^\alpha]$	$D, D^*, f, \text{DDC}, \alpha$	For IVIM: NLLS, SM, OSM, NNLS, delta. For DKI and SE: LMa	NLLS: $D: 0.01-3.5$ (ss 0.1) $\mu\text{m}^2/\text{ms}$ $D^*: 0.1-28.0$ (ss 1.0) $\mu\text{m}^2/\text{ms}$ $f: 0.001-0.25$ (ss 0.01) SM, OSM: $D: 0.01-3.5$ (ss 0.1) $\mu\text{m}^2/\text{ms}$ $D^*: 0.1-25.0$ (ss 1.0) $\mu\text{m}^2/\text{ms}$ $f: 0.001-0.25$ (ss 0.01) NNLS: $D: 0.01-4.0$ (ss 0.02) $\mu\text{m}^2/\text{ms}$ $D^*: 1-9.0$ (ss 0.2) $\mu\text{m}^2/\text{ms}$ $f: 0.0-1.0$ delta: $D: 0.01-2.0$ (ss 0.02) $\mu\text{m}^2/\text{ms}$ $f: 0.001-0.1$	Multiple initialization values
Mazaheri <i>et al.</i> ⁴⁷	3.0	3000–4000/78.2–80.4	SS-EPI	0, 600, 800, 1000, 1200, 1400, 1800, 2000	BE: $S(b)/S_0 = f_{\text{slow}} \cdot \exp(-b \cdot D_{\text{slow}}) + f_{\text{fast}} \cdot \exp(-b \cdot D_{\text{fast}})$ DKI: $S(b)/S_0 = \exp(-b \cdot D_K + b^2 \cdot D_K^2 \cdot K/6)$ SE: $S(b)/S_0 = \exp[-(b \cdot \text{DDC})^\alpha]$	$D_{\text{slow}}, D_{\text{fast}}, f_{\text{slow}}, f_{\text{fast}}, D_K, K, \text{DDC}, \alpha$	BeF using LMa	NR	
Langkilde <i>et al.</i> ¹⁵	3.0	4000/~100	NR	0, 250, 500, 750, 1000, 1250, 1500, 1750, 2000, 2250, 2500, 2750, 3000, 3250, 3500	BE: $S(b)/S_0 = f_{\text{slow}} \cdot \exp(-b \cdot D_{\text{slow}}) + f_{\text{fast}} \cdot \exp(-b \cdot D_{\text{fast}})$ DKI: $S(b)/S_0 = \exp(-b \cdot D_K + b^2 \cdot D_K^2 \cdot K/6)$ SE: $S(b)/S_0 = \exp[-(b \cdot \text{DDC})^\alpha]$	$D_{\text{slow}}, D_{\text{fast}}, f_{\text{slow}}, f_{\text{fast}}, D_K, K, \text{DDC}, \alpha$	BeF using LMa	NR	

Table 3. Imaging characteristics. Abbreviations: IVIM, Intravoxel Incoherent Motion Model; BE, Biexponential Model; DKI, Diffusion Kurtosis Imaging; SE, Stretched Exponential Model; D, molecular diffusion coefficient; D^* , pseudo-diffusion coefficient; f , perfusion fraction; D_{slow} , slow diffusion coefficient; D_{fast} , fast diffusion coefficient; f_{slow} , amplitude of slow diffusion coefficient; f_{fast} , amplitude of fast diffusion coefficient; D_K , diffusion coefficient corrected for kurtosis; K , kurtosis coefficient; DDC, distributed diffusion coefficient; α , heterogeneity index; ADC, apparent diffusion coefficient; FS, field strength; T, Tesla; TR, Repetition Time; TE, Echo Time; ms, milliseconds; Seq., diffusion sequence; HV, healthy volunteers; SS-EPI, Single-Shot Echo-Planar Imaging; BeF, Biexponential Fit; SLF, Simplified Linear Fit; NLRF, Non-Linear Regression Fit; LMa, Levenberg-Marquardt algorithm; Bpb, Bayesian probability-based approach; NNLS, Non Negative Least Square; SM, Segmented Method; OSM, Oversegmented Method; ss, step size; NR, not reported. Diffusion times column was not added because only 4 studies provided this acquisition parameter (see Supplementary Material –S6– for more details).

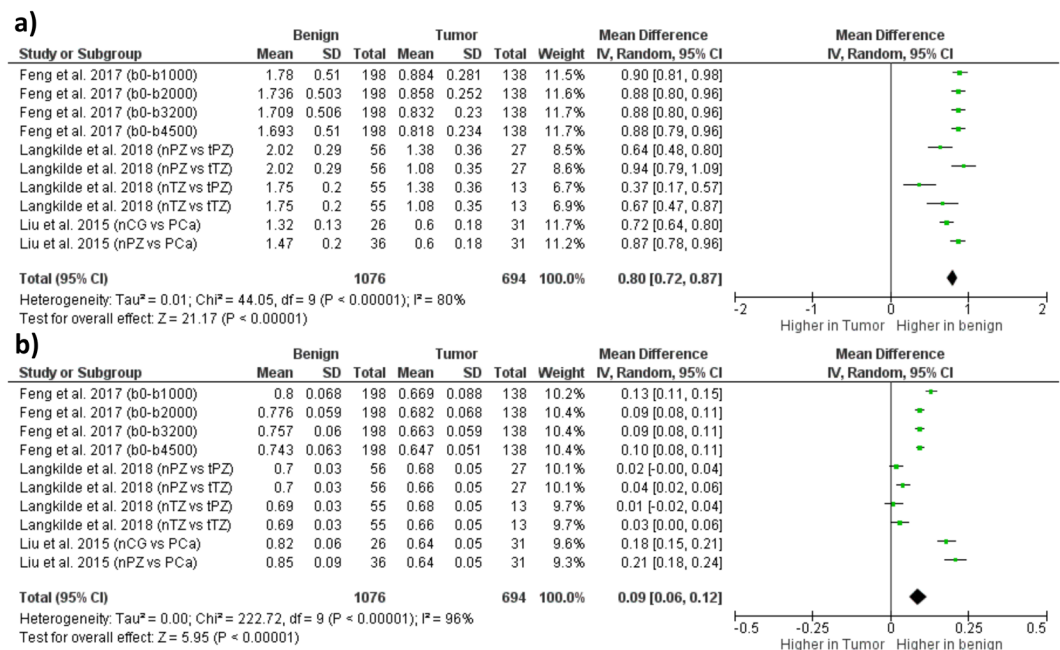


Figure 4. Forest plot showing results on Stretched Exponential Model (SE) parameters for PCa detection: (a) results on distributed diffusion coefficient (DDC) in normal tissue (NT) PCa [mean DDC \pm standard deviation (SD) $\times 10^{-3}$ mm²/s]; (b) results on heterogeneity index (α) in NT and PCa [mean $\alpha \pm$ SD]. Studies included in the meta-analysis are listed in the column “Study or Subgroup” and, in case of studies considered multiple times, the different b-value ranges and/or outcome measurements are reported in parenthesis. Abbreviations: nPZ, normal peripheral zone; tPZ, tumoral peripheral zone; nTZ, normal transitional zone; tTZ, tumoral transitional zone; nCG, normal central gland.

Publication bias. Since the number of included studies was superior to 10 only in meta-analyses based on detection and involving DKI parameters (D_K and K), publication bias was evaluated plotting funnel plots only in these cases. As shown in Supplementary Materials –S5–, all plots suggested a low risk of publication bias.

Discussions

In this study we investigated the potential value of the most commonly used Non-Gaussian DWI models for detection and characterization of prostate cancer. ME DWI has been shown to be a useful tool for PCa diagnosis, not only for detection of prostatic lesions, but also for characterization of disease aggressiveness^{8,9,16,17}. However, in the last years, due to the limited assumption at the basis of this approach, many Non-Gaussian diffusion models have been proposed to better depict diffusion signal^{19,20,23,24} and more and more studies investigated their diagnostic role and clinical value on PCa^{13,15,31,33–50,52–61}. Due to the large amount of studies and data contained therein, and the not always clear results regarding the benefits of using Non-Gaussian DWI for PCa diagnosis, we performed a systematic review followed by a meta-analysis, reviewing their parameters and trying to understand if they could provide additional information. We initially examined all found studies concerning both Non-Gaussian and Gaussian models for detection and characterization of prostate cancer from 2012 onwards. This qualitative analysis involved 29 studies: results and conclusions of selected studies varied from each other, often showing inconsistency and not a clear idea about the actual usefulness and the added value that Non-Gaussian model parameters may led to standard DWI.

In particular, among IVIM parameters, D seemed to be the most useful for PCa diagnosis, while there was no clarity on the effective significance and usefulness of the parameters related to perfusion, D^* and f . DKI parameters, K and D_K , according to a large number of examined studies were considered to be useful for PCa detection, but a clear evidence about DKI better performance had not still be reached.

Compared to those on IVIM and DKI, a smaller quantity of studies on BE and SE was present in literature, so it was even more complicated to assess the ability of these models to help in PCa detection and characterization. Among SE parameters, only DDC seemed to show significance, while conflicting views were expressed on α . About BE, there were conflicting or not significant results for D_{slow} , D_{fast} , f_{fast} , and also a not clear interpretation of their physiological basis⁴⁵.

In this scenario, it would be desirable to reach a common view on the use of Non-Gaussian DWI models in addition or in substitution of the standard DWI protocol for PCa diagnosis. With our systematic review, we summarized the available data reported in previously selected studies, and, if possible, for each Non-Gaussian model, we quantitatively evaluated the capability to differentiate between NT and PCa and between low- and intermediate/high-GS lesions. The high heterogeneity in applied functions and fitting procedures due to a not yet existing consensus on the best processing approach to fit biexponential models⁶² led us to not perform any statistical analysis on IVIM and BE models. Moreover, the small quantity of studies on SE models did not allow us to conduct statistical analysis on the power of α and DDC to classify lesions on the basis of GS.

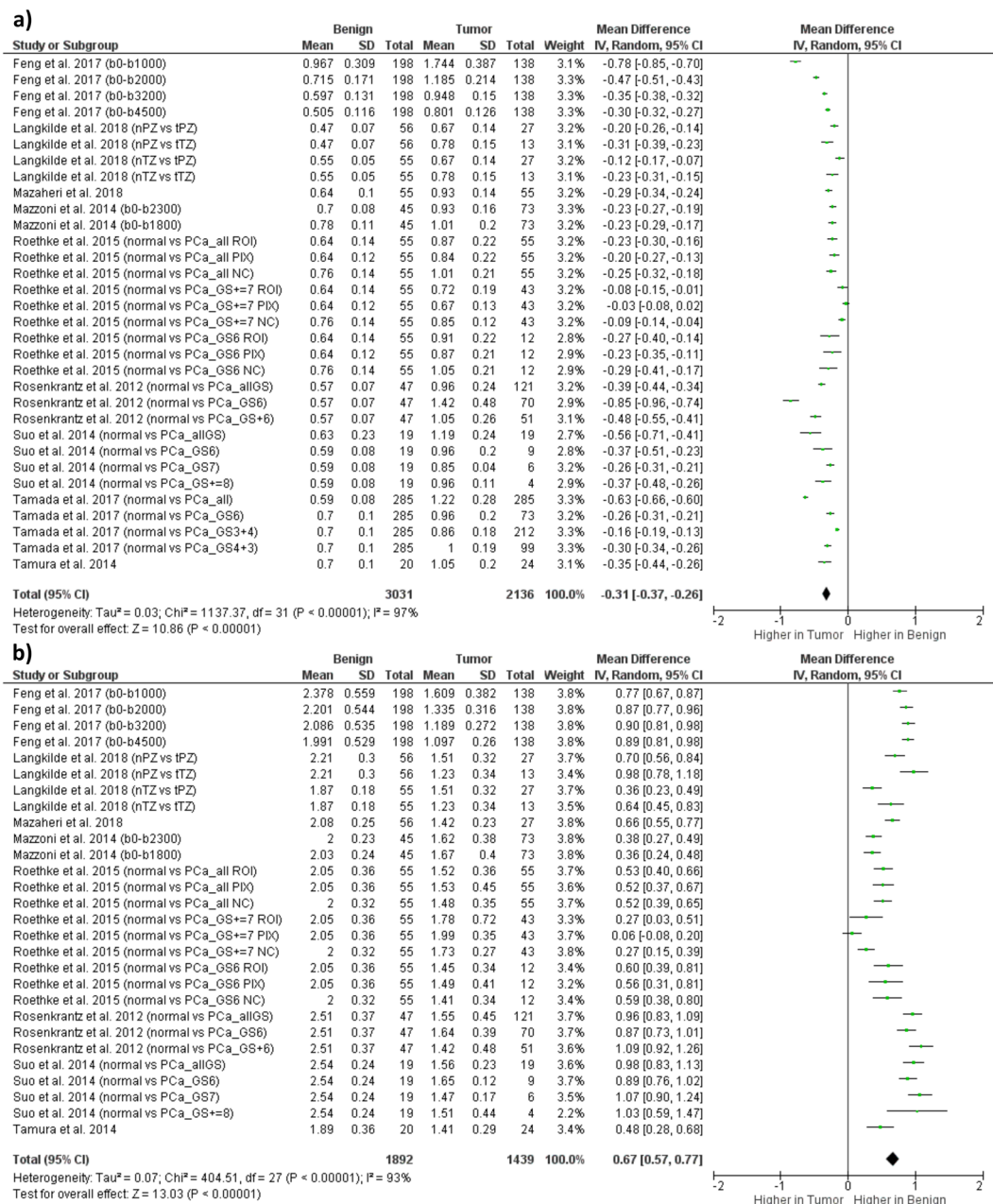


Figure 5. Forest plot showing results on Diffusion Kurtosis Imaging (DKI) model parameters for prostate cancer (PCa) detection: **(a)** results on kurtosis coefficient (K) in normal tissue (NT) and PCa [mean \pm standard deviation (SD) $\times 10^{-3}$ mm²/s]; results on diffusion coefficient corrected for kurtosis (D_K) in NT and PCa [mean \pm SD $\times 10^{-3}$ mm²/s]. Studies included in the meta-analysis are listed in the column “Study or Subgroup” and, in case of studies considered multiple times, the different b-value ranges and/or outcome measurements are reported in parenthesis. Abbreviations: nPZ, normal peripheral zone; tPZ, tumoral peripheral zone; nTZ, normal transitional zone; tTZ, tumoral transitional zone; PCa_all, all PCa lesions; PCa_GS, PCa lesions with Gleason Score equal to a certain value; PCa_GS+, PCa lesions with Gleason Score greater to a certain value; PCa_GS=, PCa lesions with Gleason Score greater or equal to a certain value; ROI, region of interest (ROI)-based fitting approach; PIX, voxel-by-voxel fitting approach; NC, ROI-based fitting approach without noise correction.

Our results showed that, concerning PCa detection, K, D_K , DDC and α showed statistical significance, allowing to distinguish prostate cancer from NT. Lower values of D_K and DDC found in PCa could be linked to the destruction of the prostatic structure-like acini together with a higher cellular and stroma density typical of PCa tissues^{49,55}. The significantly higher value of K in PCa could be associated with the increased microstructural complexity of prostate cancer^{54,55}. On the contrary, α was found to be significantly lower in PCa and this is in accordance to its association with histological heterogeneity, clearly present in prostate cancer tissue⁴⁹. It should be considered that studies involving SE were fewer than those selected for DKI and so, although the results on DDC and α showed statistical

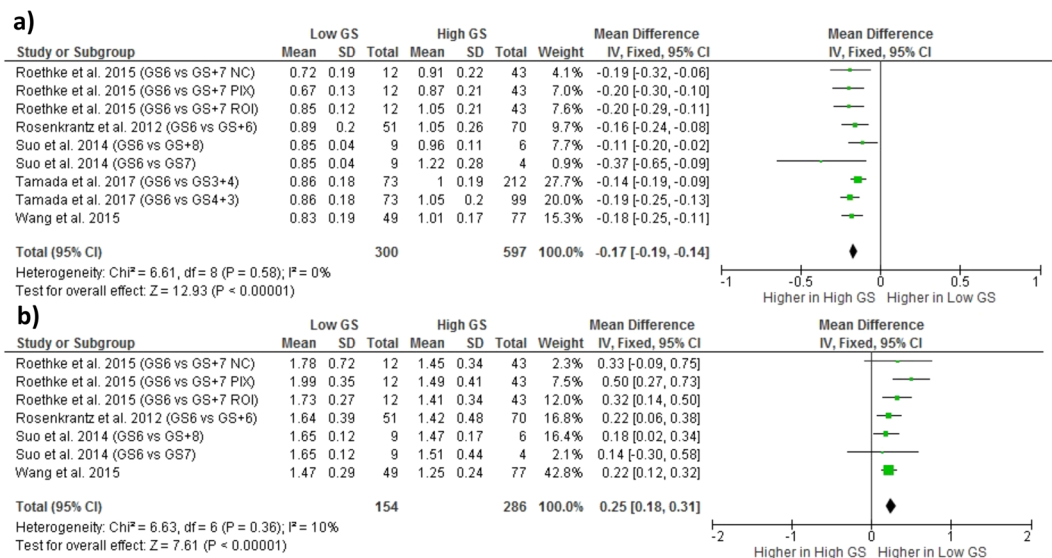


Figure 6. Forest plot showing results on DKI model parameters for prostate cancer (PCa) characterization: (a) results on diffusion kurtosis (K) in low and high Gleason Score (GS) PCa [mean \pm standard deviation (SD) $\times 10^{-3}$ mm²/s]; (b) results on diffusion coefficient corrected for kurtosis (D_K) in low and high GS PCa [mean \pm SD $\times 10^{-3}$ mm²/s]. Abbreviations: GS, PCa lesions with Gleason Score equal to a certain value; PCa_GS+, PCa lesions with Gleason Score greater to a certain value; ROI, region of interest (ROI)-based fitting approach; PIX, voxel-by-voxel fitting approach; NC, ROI-based fitting approach without noise correction.

significance, more studies would be required to validate this result. Given the high number of included studies, and the significant performance of both its parameters, DKI model claimed to be helpful for the detection of PCa.

Among outcomes regarding characterization of PCa aggressiveness, only K and D_K , showed statistical significance, differentiating lesions according to GS. The significantly greater value of D_K in intermediate/high-lesions highlighted reduced diffusion of water in these PCa than in those characterized by a GS equal or lower than 6, and it could be linked to the progressive increased cellularity of malignant lesions⁵⁴. The significantly greater value of Kurtosis (K) in intermediate/high-lesions highlight reduced diffusion of water in these lesions, and it could be linked to the progressive increased microstructural complexity of malignant lesions^{52–55}. The computed heterogeneities were lower than the ones found in the analyses on PCa detection, but it is not so reliable due to the lower number of studies involved⁶³.

The obtained high heterogeneities, both in PCa detection and characterization, may be caused by several factors that match to the limitations of the study, which are highlighted in the following subparagraph.

Limitations of the study. This meta-review suffers from several limitations which deserve to be discussed.

As concern patient characteristics, we did not consider patient age and did not draw distinction between different kind of NT ROIs (e.g. placed on prostate of healthy volunteers or on healthy prostate zones of PCa patients) or between the different examined anatomic zone (TZ, PZ, CG). Also, we have not considered the different pathologic reference standard used across studies, be it biopsy or radical prostatectomy specimens, which may lead to assign different Gleason scores for the same prostatic lesion^{64,65}.

As to characteristics of DWI protocol, because of the lack of a standard b-value range for PCa, b-values used to acquire DWI signal differed across selected studies, probably contributing to increase heterogeneity due to the supposed dependence of the computed parameters on the adopted b-values^{34,36}. Specifically for IVIM parameters, the high variability found across studies can also be explained by the effect of a non-constant Echo Time (TE)^{66,67} and of a variable number of selected b-values⁶⁸.

Moreover, the different applied diffusion time, which was found to be related to ADC and fractional anisotropy (FA) by Bourne *et al.*⁶⁹ could have influenced results, but only four of selected studies reported information on diffusion time parameters^{15,35,44,47} (see Supplementary Material –S6– for more details). Further studies aimed at finding a standard DW-MRI protocol for PCa are required.

It could be also interesting to explore the huge sources of heterogeneity and their effects on parameters performances in PCa diagnosis using a meta-regression and subgroup analyses, in order to perform a diagnostic test accuracy (DTA) meta-analysis and obtain quantitative comparison of each Non-Gaussian models with mono-exponential model, as done, for example, by Si *et al.*⁷⁰ for DKI. However, selected studies, especially those covering PCa characterization and BE or SE, did not reported sufficient data to build contingency tables, calculate quantitative measures and construct summary ROC curves. Moreover, too few studies on PCa characterization provided correlation coefficient values between parameters and GS to perform a correlation meta-analysis.

With respect to data analysis characteristics, we have not made any differences between 2D and 3D ROIs. In addition, we included studies that used both ROI-based parameters extraction approach (*i.e.*, fitting procedure over the mean signal intensity values in previously traced ROIs) and voxel-based approach (*i.e.*, voxel-by-voxel

fitting procedure and subsequent ROI measurement on the resulting parametric map⁵⁹). Further, there was neither an established fitting procedure to obtain parameters belonging to the same Non-Gaussian model and this made it impossible to perform any statistical analysis on certain models such as IVIM and BE. Larger studies are required to find a standard fitting function/procedure.

Moreover, it is not to be neglected the lack of access to data in the selected studies, which prevents the external validation of obtained results by independent researchers⁷¹. Only Feng *et al.*³⁶ took a step toward data sharing, providing, in the Supporting Material of their work, a full table including the estimated DWI parameters for each patient included in the study.

Data availability issue is gaining an increasing importance in research and in particular in MR community⁷². Some goals of data sharing should be the creation of reference findings/algorithms/implementations that can be used for comparison when publishing new methods or the possibility to compare individual results, for example in terms of consistency, computation time and hardware/programming language requirements. To our knowledge, among research teams dealing with prostate cancer diagnosis, only the groups of Chaddad *et al.*⁷³ and Jambor *et al.*⁷⁴ make their data fully available. With respect to our meta-analysis, the access to external data would have allowed not only to go beyond the traditional meta-analysis of intervention venturing into more complex meta-analytic approaches⁷⁵, but it would have also helped to shed light in the issue of standardization of DWI protocols and fitting functions/procedures for Non-Gaussian DWI parameters. The absence of standardization directly affects repeatability and reproducibility for DWI parameters that are fundamental to allow a fair and robust comparison among studies, leading to a more powerful meta-analysis⁷¹. In addition, repeatability and reproducibility are also essential in clinical practice for a correct treatment planning, response monitoring and to use DWI as clinical biomarker for cancer diagnosis^{76–82}. To our knowledge, there are only few studies on Non-Gaussian DWI models for PCa diagnosis approaching these issues. Merisaari *et al.*⁴⁴ evaluated five different IVIM fitting methods using a population of 81 patients who underwent DWI examination twice, revealing poor reproducibility of IVIM parameters. Jambor *et al.*⁴⁶ addressed for repeatability of ADC, SE, DKI and BE parameters involving a smaller patient population of 24 patients who underwent four DWI examinations, showing low repeatability for BE fitted parameters and a higher and similar repeatability among ADC, SE and DKI fitted parameters. To our knowledge, current literature lacks of studies concerning reproducibility of Non-Gaussian DWI parameters for PCa diagnosis across different centers, scanners, and readers. This is a direct consequence of the above-mentioned lack of standardization in DWI protocol, fitting functions/procedures, and also automatic tools for ROI placement⁴⁴.

Nevertheless, the strength of this work was that we assessed not only tumor detection, as in other reviews⁷⁰, but also PCa characterization in terms of aggressiveness. In addition, multiple Non-Gaussian models were considered.

Conclusions

The role of Non-Gaussian diffusion models in the detection and characterization of PCa aggressiveness and, consequently, in making clinical decisions, remains questionable. In this context, this article, highlighting the challenges that have emerged when comparing studies between each other, may serve as a starting point for future studies evaluating Non-Gaussian DWI performances to give a more precise biophysical interpretation of their parameters with the objective of identifying a standardized DW-MRI protocol in PCa diagnosis.

Received: 23 May 2019; Accepted: 28 October 2019;

Published online: 14 November 2019

References

- Bray, F. *et al.* Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA. Cancer J. Clin.*, <https://doi.org/10.3322/caac.21492> (2018).
- Gleason, D. F. Classification of prostatic carcinomas. *Cancer Chemother. reports* (1966).
- Delahunt, B., Miller, R. J., Srigley, J. R., Evans, A. J. & Samarutunga, H. Gleason grading: Past, present and future. *Histopathology* <https://doi.org/10.1111/j.1365-2559.2011.04003.x> (2012).
- Hegde, J. V. *et al.* Multiparametric MRI of prostate cancer: An update on state-of-the-art techniques and their performance in detecting and localizing prostate cancer. *Journal of Magnetic Resonance Imaging*, <https://doi.org/10.1002/jmri.23860> (2013).
- De Rooij, M., Hamoen, E. H. J., Fütterer, J. J., Barentsz, J. O. & Rovers, M. M. Accuracy of multiparametric MRI for prostate cancer detection: A meta-analysis. *Am. J. Roentgenol.* **202**, 343–351 (2014).
- Fütterer, J. J. *et al.* Can Clinically Significant Prostate Cancer Be Detected with Multiparametric Magnetic Resonance Imaging? A Systematic Review of the Literature. *Eur. Urol.* **68**, 1045–1053 (2015).
- Weinreb, J. C. *et al.* PI-RADS Prostate Imaging - Reporting and Data System: 2015, Version 2. *Eur. Urol.* <https://doi.org/10.1016/j.eururo.2015.08.052> (2016).
- Tan, C. H., Wei, W., Johnson, V. & Kundra, V. Diffusion-weighted MRI in the detection of prostate cancer: Meta-analysis. *Am. J. Roentgenol.* **199**, 822–829 (2012).
- Itou, Y., Nakanishi, K., Narumi, Y., Nishizawa, Y. & Tsukuma, H. Clinical utility of apparent diffusion coefficient (ADC) values in patients with prostate cancer: Can ADC values contribute to assess the aggressiveness of prostate cancer? *J. Magn. Reson. Imaging* **33**, 167–172 (2011).
- Falletti, R. *et al.* Can DW-MRI, with its ADC values, be a reliable predictor of biopsy outcome in patients with suspected prostate cancer? *Abdom. Radiol.* **41**, 926–933 (2016).
- deSouza, N. M. *et al.* Diffusion-weighted magnetic resonance imaging: a potential non-invasive marker of tumour aggressiveness in localized prostate cancer. *Clin. Radiol.* **63**, 774–782 (2008).
- Sato, C. *et al.* Differentiation of noncancerous tissue and cancer lesion by apparent diffusion coefficient values in transition and peripheral zones of the prostate. *J. Magn. Reson. Imaging* **21**, 258–262 (2005).
- Maurer, M. H. & Heverhagen, J. T. Diffusion weighted imaging of the prostate—principles, application, and advances. *Transl. Androl. Urol.* **6**, 490–498 (2017).
- Tamada, T., Sone, T., Jo, Y., Yamamoto, A. & Ito, K. Diffusion-weighted MRI and its role in prostate cancer. *NMR Biomed.* **27**, 25–38 (2014).
- Langkilde, F. *et al.* Evaluation of Fitting Models for Prostate Tissue Characterization Using Extended-Range b-Factor Diffusion-Weighted Imaging. **2358**, 2346–2358 (2018).

16. Kitajima, K. *et al.* Clinical utility of apparent diffusion coefficient values obtained using high b-value when diagnosing prostate cancer using 3 tesla MRI: Comparison between ultra-high b-value (2000 s/mm²) and standard high b-value (1000 s/mm²). *J. Magn. Reson. Imaging* **36**, 198–205 (2012).
17. Hambrook, T. *et al.* Relationship between Apparent Diffusion Coefficients at 3.0-T MR Imaging and Gleason Grade in Peripheral Zone Prostate Cancer. *Radiology* **259**, 453–461 (2011).
18. Le Bihan, D. Apparent Diffusion Coefficient and Beyond: What Diffusion MR Imaging Can Tell Us about Tissue Structure. *Radiology*, <https://doi.org/10.1148/radiol.13130420> (2013).
19. Le Bihan, D. *et al.* Separation of diffusion and perfusion in intravoxel incoherent motion MR imaging. *Radiology*, <https://doi.org/10.1148/radiology.168.2.3393671> (1988).
20. Mulkern, R. V. *et al.* Multi-component apparent diffusion coefficients in human brain. *NMR Biomed.*, 10.1002/(SICI)1099-1492(199902)12:1<51::AID-NBM546>3.0.CO;2-E (1999).
21. Mulkern, R. V. *et al.* Biexponential characterization of prostate tissue water diffusion decay curves over an extended b-factor range. *Magn. Reson. Imaging* **24**, 563–568 (2006).
22. Le Bihan, D. The 'wet mind': Water and functional neuroimaging. *Physics in Medicine and Biology*, <https://doi.org/10.1088/0031-9155/52/7/R02> (2007).
23. Bennett, K. M. *et al.* Characterization of continuously distributed cortical water diffusion rates with a stretched-exponential model. *Magn. Reson. Med.* **50**, 727–734 (2003).
24. Jensen, J. H., Helper, J. A., Ramani, A., Lu, H. & Kaczynski, K. Diffusional kurtosis imaging: The quantification of non-Gaussian water diffusion by means of magnetic resonance imaging. *Magn. Reson. Med.*, <https://doi.org/10.1002/mrm.20508> (2005).
25. Moher, D. *et al.* Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement (Chinese edition). *J. Chinese Integr. Med.* **7**, 889–896 (2009).
26. Whiting, P. F. *et al.* QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann. Intern. Med.*, <https://doi.org/10.7326/0003-4819-155-8-201110180-00009> (2011).
27. Higgins, J. Green, S. & Eds Cochrane handbook for systematic reviews of interventions. Version 5.1.0. Cochrane Collab. (2011).
28. Dalton, J. E., Bolen, S. D. & Mascha, E. J. Publication Bias: The Elephant in the Review. *Anesth. Analg.* **123**, 812–813 (2016).
29. Higgins, J. P. T., Thompson, S. G., Deeks, J. J. & Altman, D. G. Measuring inconsistency in meta-analyses. *BMJ*, <https://doi.org/10.1136/bmj.327.7414.557> (2003).
30. Bown, M. J. & Sutton, A. J. Quality control in systematic reviews and meta-analyses. *European Journal of Vascular and Endovascular Surgery*, <https://doi.org/10.1016/j.ejvs.2010.07.011> (2010).
31. Shinmoto, H. *et al.* An intravoxel incoherent motion diffusion-weighted imaging study of prostate cancer. *Am. J. Roentgenol.* **199**, 496–500 (2012).
32. Martin, T. *et al.* Detection of prostate cancer using IVIM model. 1–16 (2014).
33. Quentin, M. *et al.* Comparison of different mathematical models of diffusion-weighted prostate MR imaging. *Magn. Reson. Imaging* **30**, 1468–1474 (2012).
34. Mazzoni, L. N. *et al.* Diffusion-weighted signal models in healthy and cancerous peripheral prostate tissues: Comparison of outcomes obtained at different b-values. *J. Magn. Reson. Imaging* **39**, 512–518 (2014).
35. Ueda, Y. *et al.* Triexponential function analysis of diffusion-weighted MRI for diagnosing prostate cancer. *J. Magn. Reson. Imaging* **43**, 138–148 (2016).
36. Feng, Z. *et al.* Evaluation of different mathematical models and different b-value ranges of diffusionweighted imaging in peripheral zone prostate cancer detection using b-value up to 4500 s/mm². *PLoS One* **12**, 1–15 (2017).
37. Zhang, Y. D. *et al.* The Histogram Analysis of Diffusion-Weighted Intravoxel Incoherent Motion (IVIM) Imaging for Differentiating the Gleason grade of Prostate Cancer. *Eur. Radiol.* **25**, 994–1004 (2015).
38. Yang, D. M. *et al.* Prostate cancer: Correlation of intravoxel incoherent motion MR parameters with Gleason score. *Clin. Imaging* **40**, 445–450 (2016).
39. Barbieri, S., Brönnimann, M., Boxler, S., Vermathen, P. & Thoeny, H. C. Differentiation of prostate cancer lesions with high and with low Gleason score by diffusion-weighted MRI. *Eur. Radiol.* **27**, 1547–1555 (2017).
40. Bao, J. *et al.* Differentiation of prostate cancer lesions in the Transition Zone by diffusion-weighted MRI. *Eur. J. Radiol. Open* **4**, 123–128 (2017).
41. Kuru, T. H. *et al.* Intravoxel incoherent motion (IVIM) diffusion imaging in prostate cancer - What does it add? *J. Comput. Assist. Tomogr.* **38**, 558–564 (2014).
42. Valerio, M. *et al.* 3T multiparametric MRI of the prostate: Does intravoxel incoherent motion diffusion imaging have a role in the detection and stratification of prostate cancer in the peripheral zone? *Eur. J. Radiol.* **85**, 790–794 (2016).
43. Pesapane, F. *et al.* Intravoxel Incoherent Motion (IVIM) Diffusion Weighted Imaging (DWI) in the Periferic Prostate Cancer Detection and Stratification. *Med. Oncol.* **34**, 1–9 (2017).
44. Merisaari, H. *et al.* Fitting Methods for Intravoxel Incoherent Motion Imaging of Prostate Cancer on Region of Interest Level: Repeatability and Gleason Score Prediction. **1264**, 1249–1264 (2017).
45. Liu, X., Peng, W., Zhou, L. & Wang, H. Biexponential apparent diffusion coefficients values in the prostate: Comparison among normal tissue, prostate cancer, benign prostatic hyperplasia and prostatitis. *Korean J. Radiol.* **14**, 222–232 (2013).
46. Jambor, I. *et al.* 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.* **73**, 1988–1998 (2015).
47. Mazaheri, Y., Hötker, A. M., Shukla-Dave, A., Akin, O. & Hricak, H. Model selection for high b-value diffusion-weighted MRI of the prostate. *Magn. Reson. Imaging* **46**, 21–27 (2018).
48. Toivonen, J. *et al.* Mathematical models for diffusion-weighted imaging of prostate cancer using b values up to 2000 s/mm²: Correlation with Gleason score and repeatability of region of interest analysis. *Magn. Reson. Med.* **74**, 1116–1124 (2015).
49. Liu, X., Zhou, L., Peng, W., Wang, H. & Zhang, Y. Comparison of stretched-Exponential and monoexponential model diffusion-Weighted imaging in prostate cancer and normal tissues. *J. Magn. Reson. Imaging* **42**, 1078–1085 (2015).
50. Liu, W. *et al.* Histogram Analysis of Stretched- Exponential and Monoexponential Diffusion-Weighted Imaging Models for Distinguishing Low and Intermediate/High Gleason Scores in Prostate Carcinoma. 491–498, <https://doi.org/10.1002/jmri.25958> (2018).
51. Tamura, C. *et al.* Diffusion kurtosis imaging study of prostate cancer: Preliminary findings. *J. Magn. Reson. Imaging* **40**, 723–729 (2014).
52. Wang, Q. *et al.* Histogram analysis of diffusion kurtosis magnetic resonance imaging in differentiation of pathologic Gleason grade of prostate cancer. *Urol. Oncol. Semin. Orig. Investig.* **33**, 337.e15–337.e24 (2015).
53. Wang, X. *et al.* Diffusion Kurtosis Imaging Combined With DWI at 3-T MRI for Detection and Assessment of Aggressiveness of Prostate Cancer. 1–8 (2018).
54. Rosenkrantz, A. B. *et al.* Prostate Cancer: Feasibility and Preliminary Experience of a Diffusional Kurtosis Model for Detection and Assessment of Aggressiveness of Peripheral Zone Cancer. *Radiology*, <https://doi.org/10.1148/radiol.12112290> (2012).
55. Suo, S. *et al.* Non-Gaussian water diffusion kurtosis imaging of prostate cancer. *Magn. Reson. Imaging* **32**, 421–427 (2014).
56. Roethke, M. C. *et al.* Evaluation of Diffusion Kurtosis Imaging Versus Standard Diffusion Imaging for Detection and Grading of Peripheral Zone Prostate Cancer. *Invest. Radiol.* **50**, 483–489 (2015).
57. Tamada, T. *et al.* Prostate Cancer: Diffusion-weighted MR Imaging for Detection and Assessment of Aggressiveness—Comparison between Conventional and Kurtosis Models. *Radiology* **284**, 100–108 (2017).

58. Barrett, T. *et al.* Diagnostic evaluation of magnetization transfer and diffusion kurtosis imaging for prostate cancer detection in a re-biopsy population. *Eur Radiol* 3141–3150, <https://doi.org/10.1007/s00330-017-5169-1> (2017).
59. Riches, S. F., Hawtin, K., Charles-Edwards, E. M. & de Souza, N. M. 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 (2009).
60. Döpfert, J., Lemke, A., Weidner, A. & Schad, L. R. Investigation of prostate cancer using diffusion-weighted intravoxel incoherent motion imaging. *Magn. Reson. Imaging* **29**, 1053–1058 (2011).
61. Shinmoto, H. *et al.* Biexponential apparent diffusion coefficients in prostate cancer. *Magn. Reson. Imaging*, <https://doi.org/10.1016/j.mri.2008.07.008> (2009).
62. Le Bihan, D. What can we see with IVIM MRI? *Neuroimage* **187**, 56–67 (2019).
63. Ioannidis, J. P. A., Patsopoulos, N. A. & Evangelou, E. Uncertainty in heterogeneity estimates in meta-analyses. *BMJ*, <https://doi.org/10.1136/bmj.39343.408449.80> (2007).
64. Öztürk, E. *Gleason Score Correlation Between Prostate Biopsy and Radical Prostatectomy Specimens*. 1–4 (2018).
65. Khoddami, M., Khademi, Y., Aghdam, M. K. & Soltanghorae, H. *Correlation between Gleason Scores in Needle Biopsy and Corresponding Radical Prostatectomy Specimens: A Twelve-Year Review*. **11**, 120–126 (2016).
66. Feng, Z. *et al.* Effects of Echo Time on IVIM Quantification of the Normal Prostate. *Sci. Rep.* **8**, 1–8 (2018).
67. Rydhög, A. *et al.* Estimation of diffusion, perfusion and fractional volumes using a multi-compartment relaxation-compensated intravoxel incoherent motion (IVIM) signal model. *Eur. J. Radiol. Open* **6**, 198–205 (2019).
68. Ye, C. *et al.* Estimation of intravoxel incoherent motion parameters using low b-values. *PLoS One* **14**, 1–16 (2019).
69. Bourne, R. *et al.* Measurement and modeling of diffusion time dependence of apparent diffusion coefficient and fractional anisotropy in prostate tissue *ex vivo*. *NMR Biomed.* **30**, 1–10 (2017).
70. Si, Y. & Liu, R. bo. Diagnostic performance of monoexponential DWI versus diffusion kurtosis imaging in prostate cancer: A systematic review and meta-analysis. *Am. J. Roentgenol.*, <https://doi.org/10.2214/AJR.17.18934> (2018).
71. Huang, Y. & Gottardo, R. Comparability and reproducibility of biomedical data. *Brief. Bioinform.* **14**, 391–401 (2013).
72. Badji, A. Can you reproduce this seminal MRM paper? Participate in the reproducible research study group challenge! *ISMRM's MR Pulse Blog* Available at: <https://blog.ismrm.org/2019/04/02/ismrm-reproducible-research-study-group-2019-reproduce-a-seminal-paper-initiative> (2019).
73. Chaddad, A. *et al.* Predicting Gleason Score of Prostate Cancer Patients Using Radiomic Analysis. *Front. Oncol.* **8**, 630 (2018).
74. Jambor, I. *et al.* Validation of IMPROD biparametric MRI in men with clinically suspected prostate cancer: A prospective multi-institutional trial. *PLoS Med.* **16**, e1002813 (2019).
75. Leeflang, M. M. G., Deeks, J. J., Gatsonis, C. & Bossuyt, P. M. M. Systematic reviews of diagnostic test accuracy. *Annals of Internal Medicine*, <https://doi.org/10.7326/0003-4819-149-12-200812160-00008> (2008).
76. Winfield, J. M. *et al.* Extracranial soft-tissue tumors: Repeatability of apparent diffusion coefficient estimates from diffusion-weighted MR imaging. *Radiology* **284**, 88–99 (2017).
77. Jerome, N. P. *et al.* Repeatability of derived parameters from histograms following non-Gaussian diffusion modelling of diffusion-weighted imaging in a paediatric oncological cohort. *Eur. Radiol.* 345–353, <https://doi.org/10.1007/s00330-016-4318-2> (2017).
78. Plant, A. L. & Hanisch, R. J. Reproducibility and Replicability in Science, A Metrology Perspective 1 Introduction. 1–21 (2018).
79. Pathak, R. *et al.* A data-driven statistical model that estimates measurement uncertainty improves interpretation of ADC reproducibility: a multi-site study of liver metastases. *Sci. Rep.* 1–10, <https://doi.org/10.1038/s41598-017-14625-0> (2017).
80. Schmeel, F. C. Variability in quantitative diffusion-weighted MR imaging (DWI) across different scanners and imaging sites: is there a potential consensus that can help reducing the limits of expected bias? 2243–2245 (2019).
81. Malyarenko, D. *et al.* Multi-system repeatability and reproducibility of apparent diffusion coefficient measurement using an ice-water phantom. *J. Magn. Reson. Imaging*, <https://doi.org/10.1002/jmri.23825> (2013).
82. Malyarenko, D. I. *et al.* Multicenter Repeatability Study of a Novel Quantitative Diffusion Kurtosis Imaging Phantom. **5** (2019).

Acknowledgements

This work was supported by RC projects of the Italian MOH. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Author contributions

V.B. and S.M. designed the study, V.B. collected the data, performed the analyses and wrote the manuscript, with contributions from all authors. C.C. and M.S. contributed to critical discussion, S.M. supervised the analyses.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41598-019-53350-8>.

Correspondence and requests for materials should be addressed to C.C.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019