Magnetic Resonance Imaging 63 (2019) 205–216

Contents lists available at ScienceDirect

Magnetic Resonance Imaging

journal homepage: www.elsevier.com/locate/mri

Original contribution

Diffusion kurtosis imaging does not improve differentiation performance of breast lesions in a short clinical protocol

Theresa Palm[a,*], Evelyn Wenkel[a], Sabine Ohlmeyer[a], Rolf Janka[a], Michael Uder[a], Elisabeth Weiland[b], Sebastian Bickelhaupt[c], Mark E. Ladd[d,e], Maxim Zaitsev[f], Bernhard Hensel[g], Frederik B. Laun[a]

[a] Institute of Radiology, University Hospital Erlangen, Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), Maximiliansplatz 3, 91054 Erlangen, Germany
[b] Siemens Healthcare, Allee am Röthelheimpark 2, 91052 Erlangen, Germany
[c] Department of Radiology, German Cancer Research Center, Im Neuenheimer Feld 280, 69120 Heidelberg, Germany
[d] Medical Physics in Radiology, German Cancer Research Center, Im Neuenheimer Feld 280, 69120 Heidelberg, Germany
[e] Faculty of Physics and Astronomy and Faculty of Medicine, University of Heidelberg, Im Neuenheimer Feld 226, 69120 Heidelberg, Germany
[f] Department of Diagnostic Radiology, Medical Physics, University Hospital Freiburg, Breisacher Straße 60a, 79106 Freiburg, Germany
[g] Max-Schaldach-Stiftungsprofessur, Center for Medical Physics and Engineering, Friedrich-Alexander-Universität Erlangen-Nürnberg, Henkestraße 91, 91052 Erlangen, Germany

ARTICLE INFO

Keywords:
Diffusion magnetic resonance imaging
Breast neoplasms
Neoplasms by histologic type
Diagnostic imaging
Diffusion kurtosis imaging

ABSTRACT

Background: Diffusion weighted magnetic resonance imaging (DWI) is known to differentiate between malignant and benign lesions via the apparent diffusion coefficient (ADC). Here, the value of diffusion kurtosis imaging (DKI) for differentiation and further characterization of benign and malignant breast lesions and their subtypes in a clinically feasible protocol is investigated.

Material and methods: This study included 85 patients (with 68 malignant and 73 benign lesions) who underwent 3 T breast DWI using three *b* values (50, 750, 1500 s/mm^2), with a total measurement time < 5 min. ADC maps were calculated from *b* values 50, 750 s/mm^2. The diffusion kurtosis model was fitted to the diffusion weighted images, yielding in each lesion the average kurtosis-corrected diffusion coefficient D_K and mean kurtosis K. Histopathology was obtained of radiologically suspicious lesions; follow-up scans were used as a standard of reference for benign appearing lesions. Receiver operating characteristic curves were used to evaluate the parameters' diagnostic performance for differentiation of lesion types and grades. The difference in diffusion parameters between subgroups was analysed statistically using the Wilcoxon rank sum test and Kruskal-Wallis test, applying a Bonferroni correction for multiple testing where necessary.

Results: ADC, D_K and K showed significant differences between malignant and benign lesions ($p < 10^{-5}$). All parameters had similar areas under the curve (AUC) (ADC: 0.92, D_K: 0.91, K: 0.89) for differentiation of malignant and benign lesions. Sensitivity was highest for ADC (ADC: 0.96, D_K: 0.94, K: 0.93), as well as specificity (ADC: 0.85, D_K: 0.82, K: 0.82). ADC and D_K showed significant differences between tumor histologic grades ($p = 6.8{\cdot}10^{-4}$, $p = 6.6{\cdot}10^{-5}$, respectively), whereas K did not ($p = 0.99$). All three parameters differed significantly between subtypes of benign lesions (ADC: $p < 10^{-5}$, D_K: $p < 10^{-5}$, K: $p = 4.1{\cdot}10^{-4}$), but not between subtypes of malignant lesions (ADC: $p = 0.21$, D_K: $p = 0.25$, K: $p = 0.08$).

Conclusion: DKI parameters and conventional ADC can differentiate between malignant and benign lesions.

Abbreviations: ADC, apparent diffusion coefficient; AUC, area under the curve; DCE, dynamic contrast-enhanced; DCIS, ductal carcinoma in situ; D_K, kurtosis-corrected diffusion coefficient; DKI, diffusion kurtosis imaging; DWI, diffusion weighted magnetic resonance imaging; FA, fibroadenoma; ILC, invasive lobular carcinoma; IntrCystPap, intracystic papillary carcinoma; IQR, interquartile range; *K*, kurtosis; MRI, magnetic resonance imaging; Muc, mucinous carcinoma; NST, invasive carcinoma of no special type; ROC, receiver operating characteristic

* Corresponding author at: Universitätsklinikum Erlangen, Radiologisches Institut, Maximiliansplatz 3, 91054 Erlangen, Germany.
E-mail addresses: theresa.palm@uk-erlangen.de (T. Palm), Evelyn.Wenkel@uk-erlangen.de (E. Wenkel), Sabine.Ohlmeyer@uk-erlangen.de (S. Ohlmeyer), Rolf.Janka@uk-erlangen.de (R. Janka), Michael.Uder@uk-erlangen.de (M. Uder), Elisabeth.Weiland@siemens-healthineers.com (E. Weiland), S.Bickelhaupt@Dkfz-Heidelberg.de (S. Bickelhaupt), Mark.Ladd@dkfz-heidelberg.de (M.E. Ladd), Maxim.Zaitsev@uniklinik-freiburg.de (M. Zaitsev), Bernhard.Hensel@fau.de (B. Hensel), Frederik.Laun@uk-erlangen.de (F.B. Laun).

https://doi.org/10.1016/j.mri.2019.08.007
Received 21 May 2019; Received in revised form 26 July 2019; Accepted 15 August 2019

Differentiation performance was best for ADC. Different tumor grades were significantly different in ADC and D_K, which may have an impact on therapy planning and monitoring. In our study, K did not add value to the diagnostic performance of DWI in a clinical setting.

1. Introduction

Breast cancer is the most common cancer in women [1]. For women with high risk of breast cancer, intensified early diagnosis is recommended, including annual magnetic resonance imaging (MRI) [2–4]. MRI is also used for risk assessment and treatment planning in patients with breast cancer [5–7]. Currently used MRI protocols include dynamic contrast-enhanced (DCE) T1-weighted imaging [8–10], but the ongoing discussion about gadolinium deposition in the brain [11–14] stresses the need for alternative contrast-agent-free techniques.

Diffusion weighted imaging (DWI) is such a technique. It allows measurement of water molecular displacements, which can be used for detection and classification of breast lesions [15,16]. Several studies showed that low apparent diffusion coefficient (ADC) values indicate malignant lesions [17–21]. However, the ADC values for benign and malignant lesions overlap, which suggests that more sophisticated metrics might be better suited for data characterization [22]. The ADC does not take into account the fact that diffusion in tissue is non-Gaussian because it is restricted by cell structures. Diffusion kurtosis imaging (DKI) is one of the most prominent approaches for non-Gaussian diffusion mapping. In DKI, the excess kurtosis K describes the deviation from Gaussian diffusion and thereby the cellularity of the tissue [23]. Additionally, a kurtosis-corrected diffusion coefficient D_K can be measured [23]. Previous studies found D_K to be significantly lower in malignant breast lesions than in benign lesions and K to be significantly higher [22,24–29]. Several studies reported that the diagnostic performance (AUC) for differentiation of malignant and benign lesions is higher for DKI parameters than for conventional ADC [25,27,30,31].

Further imaging-based classification of breast lesions could accelerate patient care, e.g. by differentiation between subgroups of lesions like histologic type and grade. So far, only few studies on breast DKI have reported on subgroup differentiation: Suo et al. found D_K to be significantly lower and K to be higher in invasive cancer than in in situ cancer ($p < 0.05$) [24]. Contrary to this, Iima et al. reported no statistically significant difference in D_K and K between molecular subtypes of breast cancer [25]. Furthermore, Sun et al. and Huang et al. reported that DKI parameters correlate with lesion grading [27,30]. Investigating benign lesions, Nogueira et al. found that only K differed significantly between two subtypes (fibroadenomas and fibrocystic changes), but D_K did not [28].

The aim of this work was to apply clinically feasible DKI and DWI protocols with measurement time < 5 min to patients with benign and malignant breast lesions in order to further investigate the power of ADC, D_K and K for differentiation of malignant and benign breast lesions. Furthermore, it was investigated if ADC, D_K and K can differentiate between histologic subtypes of malignant and benign lesions and between tumor grades.

2. Material and methods

2.1. Patients

This study was performed in accordance with the Declaration of Helsinki and received approval by the local institutional review board. Written informed consent was obtained from the patients prior to examination. A total of 382 patients with various indications underwent breast MRI with the study protocol between April 2017 and February 2018. Indications for the MRI exam included suspicious or inconclusive sonography or mammography, screening of women with high risk for breast cancer or history of breast cancer. The patients were examined with the DKI sequence in addition to the standard clinical breast MRI protocol. Exclusion criteria for this study were: no lesion present and breast implants. Ultimately, 85 patients were included in this study. One patient was subsequently excluded from the study because of poor image quality in the DWI. In 40 patients, more than one lesion per subject was included in the study. Histopathology was used as reference standard for all radiologically suspicious lesions. It was obtained by core-needle biopsy, vacuum biopsy, and/or surgery. Lesions without histological verification were defined as benign if they were stable from previous or follow-up examinations for >12 months or if they were

Table 1
Imaging protocol.

Parameter	DWI/DKI	T2-weighted	DCE
Sequence	Single-shot echo planar	T2 STIR	T1 Dixon
Orientation	Transversal	Transversal	Transversal
Repetition time (ms)	6290–9065	3570–4060	5.97
Echo time (ms)	66	70	TE_1: 2.46, TE_2: 3.69
Voxel size (mm^3)	1.4 × 1.4 × 4.0	0.8 × 0.8 × 4.0	0.8 × 0.8 × 1.5
Fat suppression	Inversion recovery (TI = 220 ms)	Inversion recovery (TI = 230 ms)	–
Field of view (mm^2)	(350 × 218)–(430 × 268)	(340 × 340)–(399 × 399)	(360 × 360)–(379 × 379)
Matrix	128 × 128 (160 × 256)	448 × 448 (448 × 358)	448 × 448 (448 × 385)
Slice thickness (mm)	4.0	4.0	1.5
No. of slices	34–49	34–49	112
Multi-slice mode	Interleaved	Interleaved	Sequential
Parallel imaging	GRAPPA	GRAPPA	GRAPPA
Acceleration factor	2	2	3
Bandwidth (Hz/pixel)	2298	248	800
Acquisition time (min)	3:34–4:56	3:29–4:30	6:21
$\boldsymbol{b}$ values (s/mm^2)	50, 750, 1500		
Averages	3, 8, 20		
Diffusion mode	3D diagonal		
Diffusion scheme	monopolar		

sonographically classical cysts [32]. This classification was performed by experienced radiologists as part of the clinical routine. For the carcinomas, histologic subtype and grade were determined from histopathologic reports. For one mucinous carcinoma (Muc), one invasive ductal carcinoma (NST), and three ductal carcinomas in situ (DCIS), grading was not available from the pathological report.

2.2. MRI exam

All patients were examined with a standard clinical breast MRI protocol at a 3 T MR scanner (MAGNETOM Skyra, Siemens Healthcare, Erlangen, Germany). In addition, for DWI and DKI, a non-product sequence with inversion recovery fat saturation was applied. The subjects were scanned in prone position using a dedicated 16-channel bilateral breast coil. The parameters of the sequences used are stated in Table 1.

2.3. Image analysis

All evaluated lesions were larger than 5 mm in diameter. Lesion location was double checked on the DCE image (second subtraction) for the quantitative evaluation of contrast-enhancing malignant and benign lesions. Regions of interest (ROIs) were placed manually on the b_{1500} image of the slice with the largest lesion area, excluding the most peripheral parts of the lesion and necrotic regions. All non-contrast-enhancing benign lesions (cysts and haematomas) were delineated only on the D_K map, also placing the ROI in the slice with the largest lesion area. The ROIs were checked by a radiologist with 15 years of experience in breast MRI. The mean ROI size was 60 mm^2 (range 8–1083 mm^2). Mean ADC values of the ROIs were calculated from the signal intensity at b values 50 and 750 s/mm^2. K and D_K of the ROI were calculated using three b values (50, 750, 1500 s/mm^2). A least-squares fit to the following equation was performed using self-written Matlab code (The MathWorks, Inc.):

$$\frac{S(b)}{S(0)} = \exp\left(-bD_K + \frac{1}{6}b^2D_K^2K\right),$$

where $S(b)$ stands for the signal intensity in dependency of the b value and $S(0)$ is the signal without diffusion weighting. No image filter was applied before fitting the data. The fit was performed voxel by voxel to obtain maps of D_K and K. Moreover, the fit was performed on the ROI-averaged signal for the quantitative evaluation.

2.4. Statistical analysis

ADC, D_K and K were compared between subgroups of lesions using the Wilcoxon rank sum test (in case of 2 groups) and the Kruskal-Wallis test (in case of >2 groups). Bonferroni correction was applied to correct for multiple testing. A p-value lower than 0.05 was considered to indicate a significant difference. Statistical analyses were performed in Matlab and R [33].

In order to assess the diagnostic performance of the parameters ADC, D_K and K, a ROC analysis was performed. Sensitivity and specificity were calculated based on the Youden index. The AUCs of the three parameters were compared using the method developed by DeLong et al. [34] and the p-values of the AUCs were calculated with the pROC tool in R (R Foundation for Statistical Computing) [35].

3. Results

Out of all 141 lesions, 68 were histologically categorized as malignant. The carcinomas comprised 44 NST, 17 invasive lobular carcinomas (ILC), 4 DCIS, 2 encapsulated papillary carcinomas (IntrCystPap), and 1 Muc. The 73 benign lesions were categorized as 41 cysts (including 2 complex cysts), 19 fibroadenomas (FA), 5 haematomas, 4 fibrocystic changes, 2 fat necroses, 1 papilloma, and 1 adenosis tumor.

For all lesions in this study, the ranges of quantitative values were: for malignant lesions ADC [0.44; 2.50] μm^2/ms, D_K [0.51; 2.85] μm^2/ms, K [0.15; 3.05] and for benign lesions ADC [0.62; 2.93] μm^2/ms, D_K [0.72; 4.09] μm^2/ms, K [0.13; 1.47]. Medians are listed in Table 2 in combination with 25th and 75th percentile for benign and malignant lesions. For malignant lesions, values are subclassified in grading and histological subtypes.

Representative images of two contrast-enhancing spiculated carcinomas (NST, G2) of the left breast are shown in Fig. 1. As an example for benign lesions, a fibroadenoma in the left breast of a different patient is shown in Fig. 2. In comparison to Fig. 1b, contrast is reduced between the lesion and healthy breast tissue in the DCE image (Fig. 2b). D_K is higher and K is lower than in the malignant lesions of Fig. 1. The benign lesion appears more homogeneous on the K map and more heterogeneous on the D_K map compared to the carcinoma in Fig. 1.

Fig. 3 shows boxplots of ADC values of all lesions grouped into malignant and benign lesions (Fig. 3a) and lesion subtypes (Fig. 3b–d). ADC values for benign and malignant lesions differed significantly ($p < 10^{-5}$). For different histologic grades of the malignant lesions (Fig. 3b), a significant difference in ADC between the grades was found ($p = 6.8 \cdot 10^{-4}$). Higher grades had lower ADC values. Grade 1 lesions differed significantly in ADC from higher grade lesions (G2 and G3) ($p = 8.2 \cdot 10^{-3}$). ADC values of the four subtypes of benign lesions (haematomas, cysts, FA, and others) were significantly different ($p < 10^{-5}$) (Fig. 3c). Cysts had a higher ADC than FA ($p = 4.3 \cdot 10^{-5}$). Comparing the histologic subtypes of malignant lesions (Fig. 3d), ADC

Table 2
Diffusion parameters (median, 25th percentile, 75th percentile).

Parameter	ADC in μm^2/ms	D_K in μm^2/ms	K
Malignant (n = 68)	**0.96 (0.7, 1.14)**	**1.19 (0.94, 1.40)**	**1.03 (0.88, 1.37)**
Histologic grade			
Grade 1 (n = 10)	1.08 (1.00, 1.22)	1.44 (1.30, 1.60)	1.01 (0.95, 1.52)
Grade 2 (n = 38)	0.98 (0.79, 1.14)	1.20 (0.95, 1.42)	1.04 (0.90, 1.33)
Grade 3 (n = 15)	0.77 (0.68, 0.86)	0.87 (0.76, 1.07)	1.08 (0.83, 1.51)
Histologic subtype			
Muc (n = 1)	2.50 (−)	2.85 (−)	0.33 (−)
IntrCystPap (n = 2)	1.51 (1.00, 2.03)	1.88 (1.17, 2.59)	0.79 (0.63, 0.95)
NST (n = 44)	0.96 (0.79, 1.13)	1.20 (0.90, 1.38)	1.02 (0.88, 1.22)
DCIS (n = 4)	1.02 (0.80, 1.21)	1.23 (1.03, 1.44)	1.22 (0.79, 1.66)
ILC (n = 17)	0.88 (0.66, 1.05)	1.05 (0.83, 1.35)	1.31 (0.97, 1.55)
Benign (n = 73)	**2.04 (1.60, 2.38)**	**2.37 (1.86, 2.79)**	**0.51 (0.37 0.64)**
Haematomas (n = 5)	2.78 (2.65, 2.92)	3.18 (3.06, 3.23)	0.29 (0.24, 0.34)
Cysts (n = 41)	2.22 (2.03, 2.44)	2.58 (2.31, 2.87)	0.45 (0.29, 0.54)
Fibroadenomas (n = 19)	1.61 (1.36, 1.83)	1.85 (1.62, 2.17)	0.58 (0.52, 0.90)
Other (n = 8)	1.32 (1.06, 1.52)	1.47 (1.32, 1.79)	0.67 (0.47, 1.11)

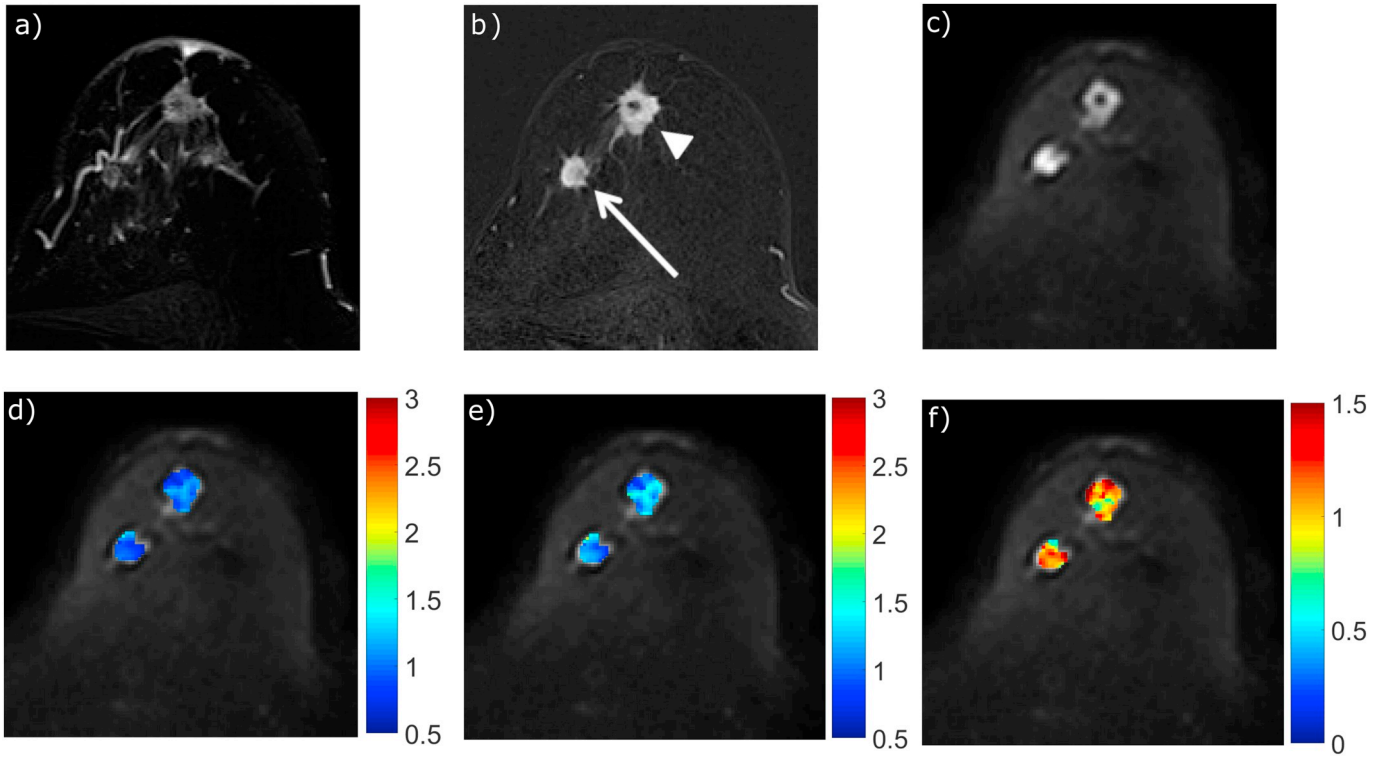

Fig. 1. Two breast cancers (NST, Grade 2) in different MR contrasts a) T2-weighted image, b) contrast-enhanced second subtraction (arrow: NST, arrowhead: NST with central necrosis), c) diffusion-weighted image, $b = 1500\,\mathrm{s/mm^2}$, d) ADC map in $\mu\mathrm{m^2/ms}$, e) D_K map in $\mu\mathrm{m^2/ms}$, f) kurtosis map (unitless). The lesion closer to the nipple (arrowhead) exhibits central necrosis. The kurtosis map is heterogeneous in both lesions, with low values in the necrotic area. D_K is higher in the central necrosis. The smaller lesion (arrow) also has low D_K values that correspond to high values on the K map.

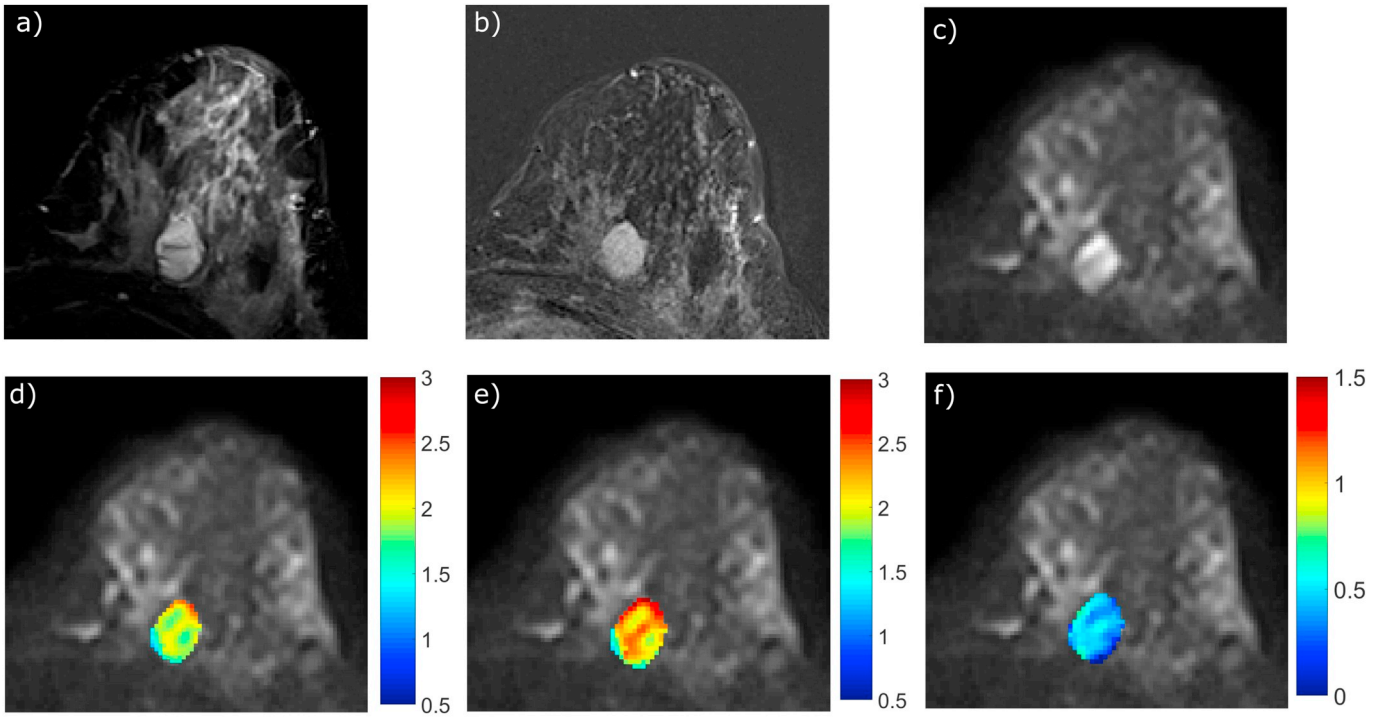

Fig. 2. Example of a benign lesion (fibroadenoma) in different MR contrasts a) T2-weighted image, b) contrast-enhanced second subtraction, c) diffusion-weighted image, $b = 1500\,\mathrm{s/mm^2}$, d) ADC map in $\mu\mathrm{m^2/ms}$, e) D_K map in $\mu\mathrm{m^2/ms}$, f) kurtosis map (unitless).

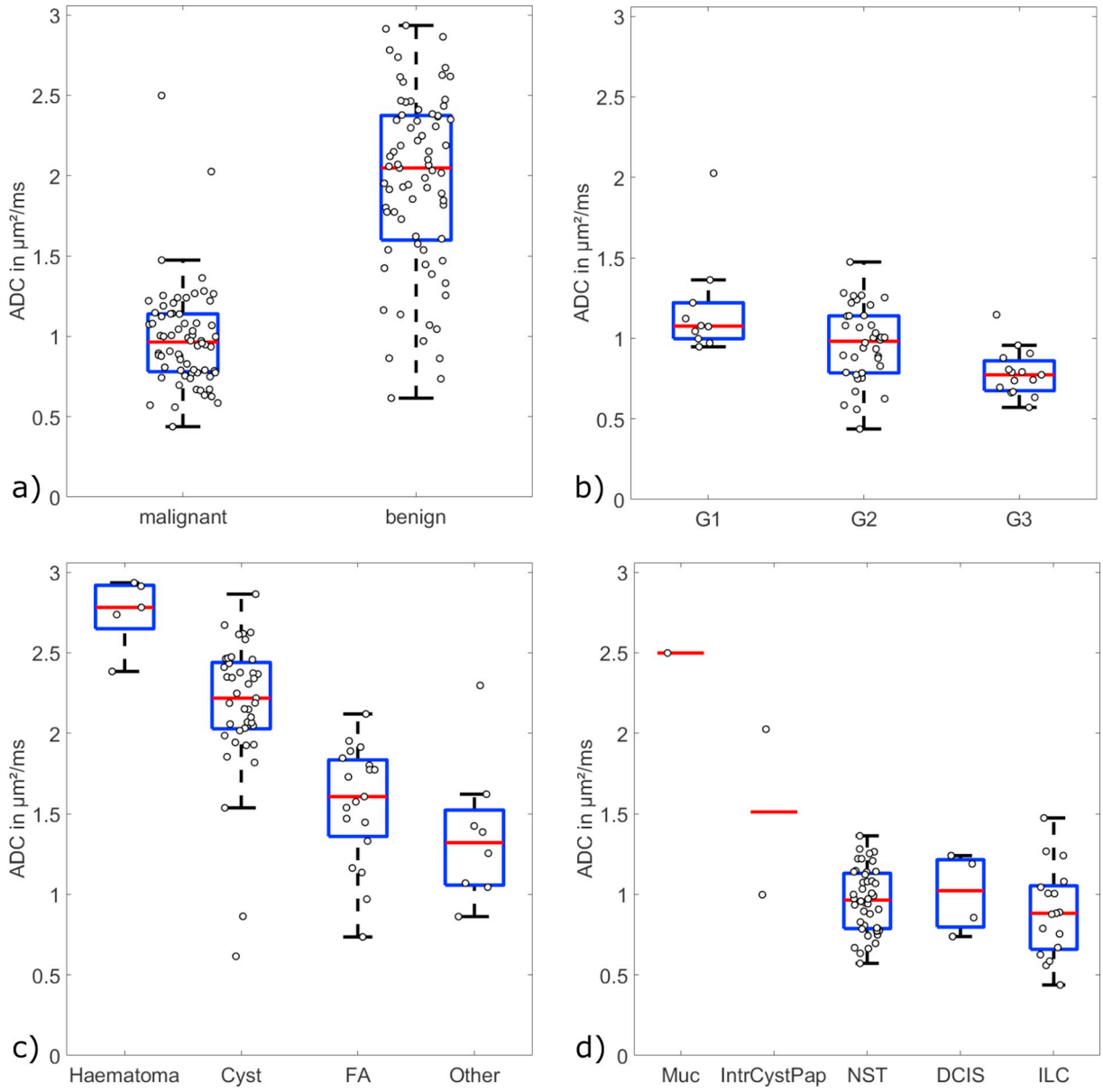

Fig. 3. Boxplots of the ADC values of lesion subgroups a) malignant and benign lesions, b) grading of malignant lesions, c) types of benign lesions, d) histologic subtypes of malignant lesions. The red line indicates the median of the data. Blue boxes show the 25th and 75th percentile; whiskers indicate minimum and maximum values except for outliers. Data points are plotted as well, also indicating the outliers. The plot shown in b) contains fewer data points than in a) because grading was not available for 1 Muc and 3 DCIS. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

ranges of DCIS, ILC, and NST overlapped. There was no significant difference between the malignant subtypes' ADC values ($p = 0.21$). The Muc had a higher ADC than all other malignant lesions.

The results for D_K values of all lesions grouped into malignant and benign lesions (Fig. 4a) and lesion subtypes (Fig. 4b–d) are similar. Apart from two outliers (1 Muc and 1 IntrCystPap), all lesions with $D_K > 1.81\,\mu m^2/ms$ were benign. D_K values for benign and malignant lesions differed significantly ($p < 10^{-5}$). With regard to the histologic grading of the malignant lesions (Fig. 4b), a significant difference in D_K between the grades was found ($p = 7.6 \cdot 10^{-5}$). Higher grades were accompanied by lower D_K values. Grade 1 lesions differed significantly in D_K from higher grade lesions (G2 and G3) ($p = 1.4{\cdot}10^{-3}$). D_K values of the four subtypes of benign lesions (haematomas, cysts, FA, and others) were significantly different ($p < 10^{-5}$) (Fig. 4c). Cysts had a higher D_K than FA ($p = 1.5{\cdot}10^{-4}$). Comparing the histologic subtypes of malignant lesions (Fig. 4d), D_K ranges of DCIS, ILC, and NST overlapped. There was no significant difference between the malignant subtypes' D_K values ($p = 0.25$). The Muc had a higher D_K than all other malignant lesions.

The kurtosis of benign lesions was found to be significantly lower than that of malignant lesions ($p < 10^{-5}$) (Fig. 5a). Unlike the diffusion coefficients ADC and D_K, K only showed a weak dependency on tumor grade (Fig. 5b). The differences in K between the tumor grades were not significant ($p = 0.99$), nor could Grade 1 lesions be differentiated from higher grade lesions ($p = 0.93$) based on K. The kurtosis values of benign lesions behaved inversely compared to their ADC and D_K values

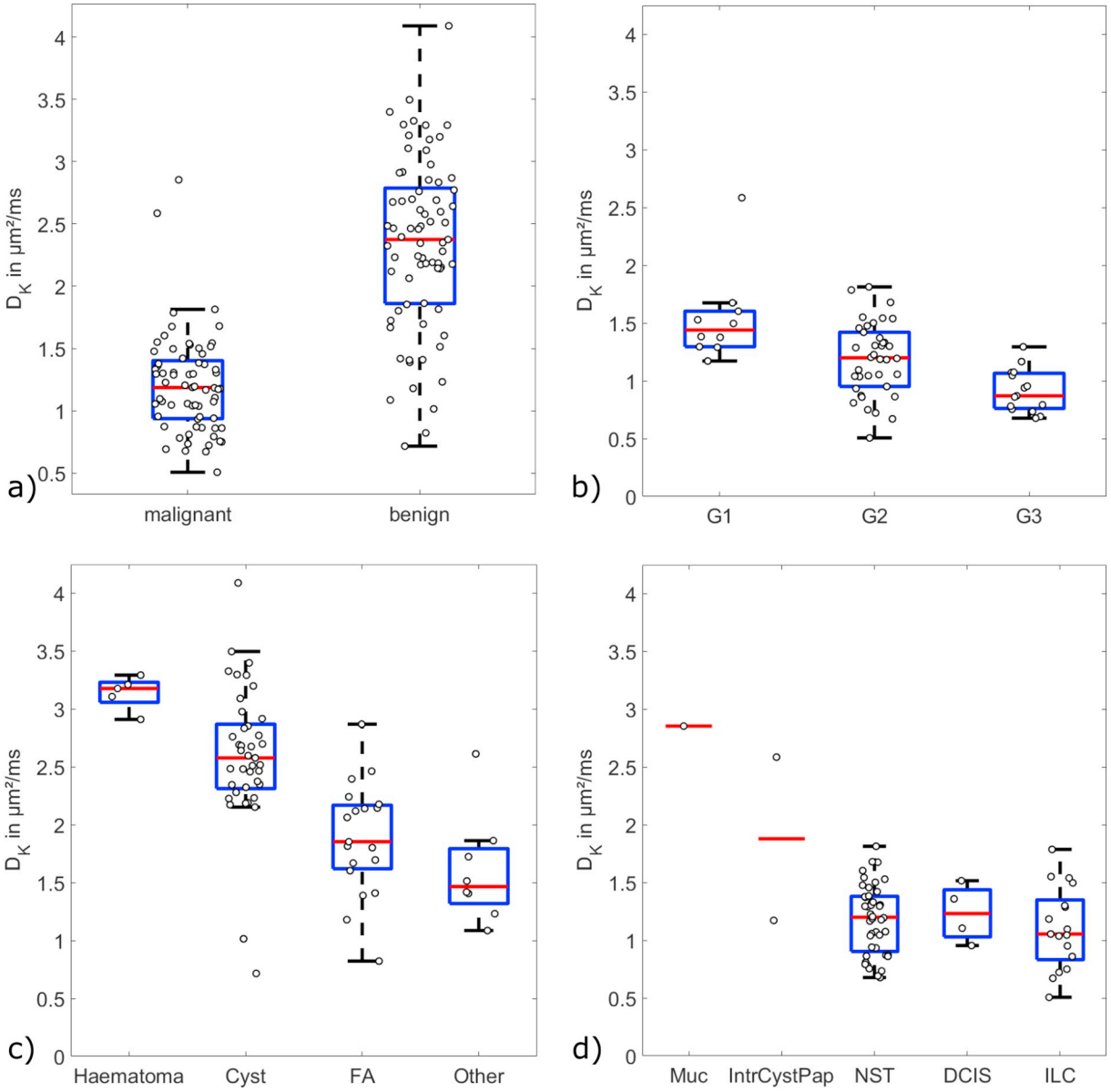

Fig. 4. Boxplots of the D_K values of lesion subgroups a) malignant and benign lesions, b) grading of malignant lesions, c) types of benign lesions, d) histologic subtypes of malignant lesions. The red line indicates the median of the data. Blue boxes show the 25th and 75th percentile; whiskers indicate minimum and maximum values except for outliers. Data points are plotted as well, also indicating the outliers. The plot shown in b) contains fewer data points than in a) because grading was not available for 1 Muc and 3 DCIS. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

(Fig. 5c). For example, cysts and haematomas had high D_K but low K values. K differed significantly between all benign subgroups ($p = 4.2{\cdot}10^{-4}$). Considering cysts and FA specifically, the difference in K was also significant ($p = 9{\cdot}10^{-3}$). Among the malignant lesions, NST had a lower median K value than ILC and DCIS, while their K value ranges overlapped. K did not differ significantly between malignant subtypes ($p = 0.08$). The single Muc had a lower kurtosis value than any other malignant lesion apart from one NST outlier.

Fig. 6 shows the ROC curves of ADC, D_K and K for differentiation of benign from malignant lesions, FA from malignancies, FA from ILC and NST, and ILC from NST. None of the AUC values differed significantly between ADC, D_K and K ($p \geq 0.24$ for all AUCs).

The performance in differentiating between malignant and benign lesions was almost the same for ADC, D_K and K (AUC = 0.92, 0.91 and 0.89, $p \geq 0.46$, sensitivity = 0.96, 0.94 and 0.93, specificity = 0.85, 0.82 and 0.82).

Looking only at FA versus malignancies, ADC and D_K slightly outperformed K in terms of AUC (0.88, 0.87 vs 0.85). Again, none of the differences in AUC were significant ($p \geq 0.45$). ADC and D_K had a higher specificity than K (0.79, 0.79 vs 0.68) and ADC and K had a higher sensitivity than D_K (0.94, 0.94 vs 0.91).

The same principal picture holds for the differentiation of FA from ILC and NST. The AUC (0.91, 0.90 vs 0.87) and specificity (0.79, 0.79 vs 0.68) were higher for ADC and D_K than for K and the sensitivity was

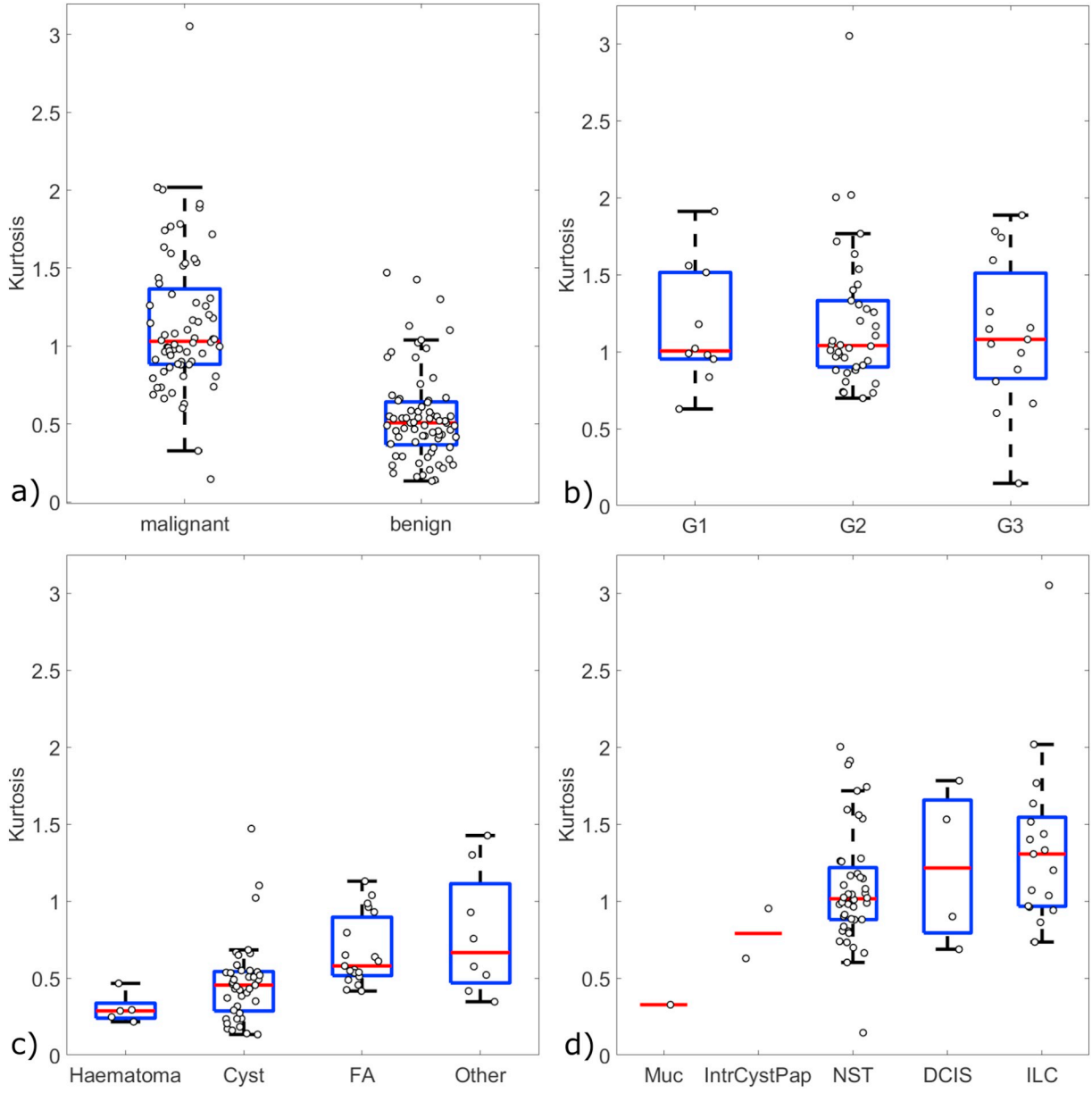

Fig. 5. Boxplots of the K values of lesion subgroups a) malignant and benign lesions, b) grading of malignant lesions, c) types of benign lesions, d) histologic subtypes of malignant lesions. The red line indicates the median of the data. Blue boxes show the 25th and 75th percentile; whiskers indicate minimum and maximum values except for outliers. Data points are plotted as well, also indicating the outliers. The plot shown in b) contains fewer data points than in a) because grading was not available for 1 Muc and 3 DCIS. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

highest for ADC and K compared to D_K (0.97, 0.97 vs 0.93). No significant differences were found between the AUCs ($p \geq 0.24$).

For distinguishing between ILC and NST, ILC was defined as the positive class in the ROC. All three diffusion parameters were close to random performance, with K better than ADC and D_K (AUC = 0.66 vs 0.59 and 0.58, sensitivity = 0.53 vs 0.59 and 0.65, specificity = 0.82 vs 0.64 and 0.55). The AUCs did not differ significantly ($p \geq 0.29$).

The optimum cut-off values for all tasks that involved differentiation between malignant and benign lesions were similar (ADC: 1.36, 1.28 and 1.28 $\mu m^2/ms$, D_K: 1.68, 1.60, and 1.60 $\mu m^2/ms$, K: 0.69, 0.66, and 0.66, for malignant vs benign, FA vs malignancies, FA vs ILC and NST, respectively). For distinguishing ILC from NST, the cut-off values for K and D_K were in a different range (ADC: 0.89 $\mu m^2/ms$, D_K: 1.18 $\mu m^2/ms$, K: 1.31, Fig. 5d).

4. Discussion

4.1. Summary and interpretation of results

All three diffusion parameters showed significant differences between malignant and benign lesions. ADC and D_K behaved similarly, which is expected as D_K is the kurtosis-corrected diffusion coefficient. The malignant outliers in ADC and D_K were one Muc and one IntrCystPap, which both have reportedly false negative ADCs [36–38].

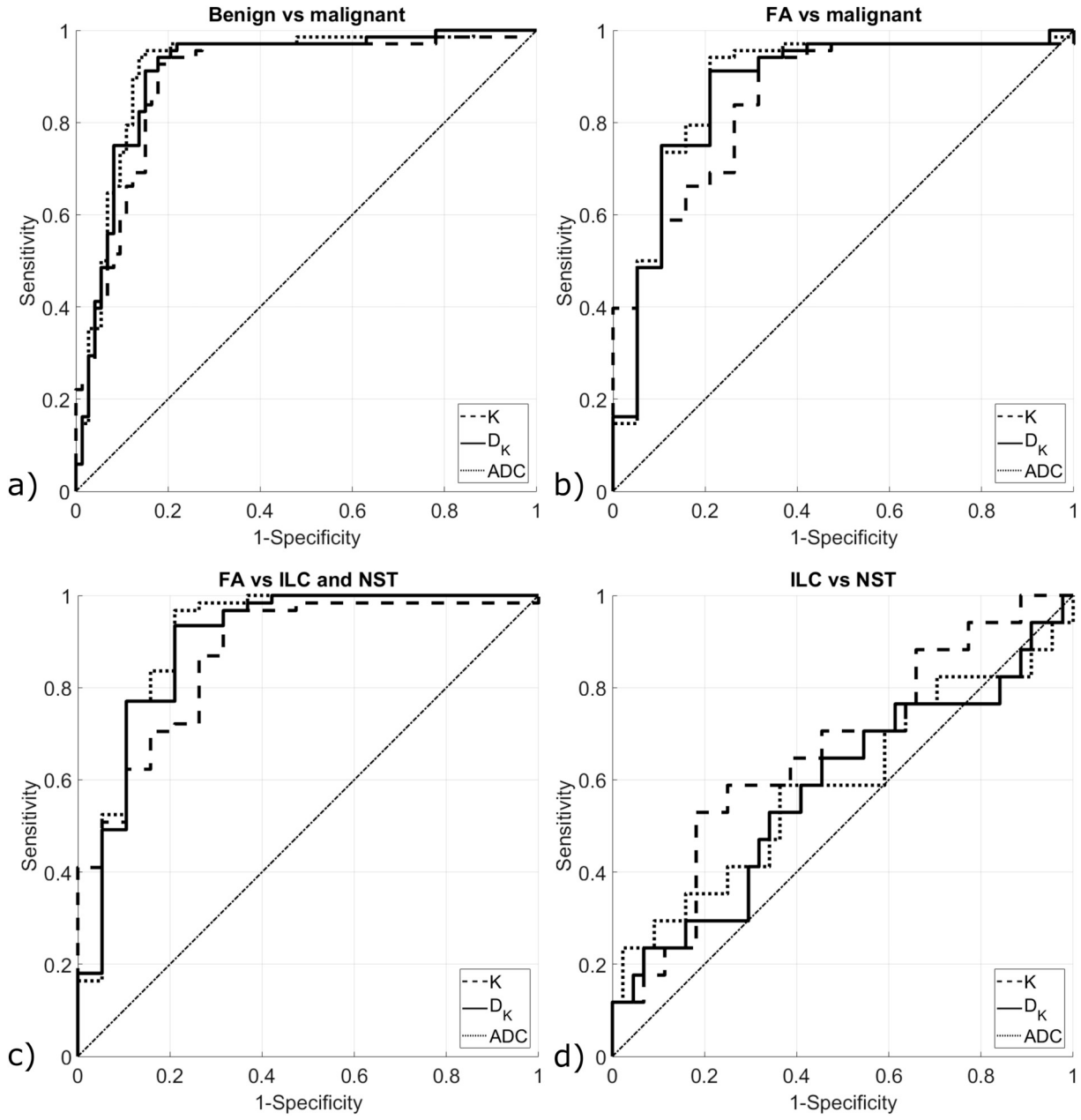

Fig. 6. ROC curves for differentiation of lesion subgroups: a) benign versus malignant lesions, b) fibroadenoma versus malignant lesions, c) fibroadenoma versus ILC and NST, d) ILC versus NST.

Both lesions had low K values, which is expected given the inverse relationship between the diffusion coefficients and K. In terms of differentiation performance (malignant vs benign), ADC had a higher AUC than D_K and K, although the difference was not significant.

ADC and D_K differed significantly between the grades of malignant lesions, while K did not. Higher-grade lesions generally had a lower diffusion coefficient and higher K values, which presumably reflects a higher cell density. Neither ADC, D_K nor K were able to differentiate between histologic subtypes of malignancies.

These results indicate that DKI parameters can differentiate between malignant and benign breast lesions. However, the conventional ADC performed better than DKI parameters in every differentiation task between malignant and benign lesion. Thus, in a clinical setup, the quantitative tissue information provided by the ADC suffices to reduce unnecessary biopsies or follow-ups for benign lesions. DKI with three b values in a short clinical protocol adds little value to the diagnostic process.

4.2. Literature comparison

Comparing our findings on the differentiation performance of DKI parameters and ADC to those reported in the literature, there is a discrepancy. One possible explanation is that the studies that reported higher AUCs for DKI parameters than for ADC used long protocols with

5 or more b values up to $b = 3000\,s/mm^2$ for DKI, while the ADC was calculated from 2 b values $\leq 1000\,s/mm^2$. They all have in common that they used more b values for DKI measurements than for DWI [25,27,30,31]. In our study, we investigated how well a DKI protocol with the minimum number of b values performs. The acquisition time was <5 min, which is lower than in [30,31] but not than in [25,27].

Iima et al. reported that DKI parameters determined from a limited MRI protocol using few b values (5, TA 3:30 min, $b_{max} = 2500\,s/mm^2$) were not statistically different from DKI parameters obtained via a protocol with many b values from the same range (16, TA: 3:55 min) [39]. In our study, DKI parameters have a lower differentiation performance than conventional ADC when using three b values $\leq 1500\,s/mm^2$. A study by Suo et al. found that using the same number and range of b values for DKI and DWI, ADC is superior to DKI parameters in terms of differentiation between malignant and benign lesions [24]. Taking all these results into consideration, it could be argued that while reducing the number of b values DKI, loss of diagnostic performance could possibly be avoided by maintaining the highest b value at $>1500\,s/mm^2$. In order to further investigate how the number and range of b values change the DKI parameters and whether these changes affect the differentiation performance in this study, an extended DKI protocol was used for additional measurement and simulations (see Appendix).

The wide spread of reported mean/median values, AUC, and cut-off values in the literature presumably originates from the variance in measurement parameters (b values, TE, B_0), patient cohorts, and evaluation methods [22,24–31]. For example, voxel-based fitting of the data was performed in most of the studies [24,26,27,29–31], whereas the others fitted the ROI mean signal intensity (our study, [22,25,28]). Some studies used 3D ROIs [24,27,29], while others evaluated only one slice (our study, [22,25,26,28,30,31]). These differences can cause discrepancies in the reported values, and stress the need for harmonization of protocols and post-processing procedures.

In agreement with this study, Sun et al. and Huang et al. found a negative correlation of D_K and histologic grade for invasive cancer [27,30]. The other correlation found in our study between tumor histologic grade and ADC is supported by findings by Rahbar et al. [40], who reported significantly higher ADC in low- and moderate-grade DCIS than in low-grade DCIS or normal tissue. Furthermore, Sun et al. and Huang et al. found K to positively correlate with grading, which was not the case in this study. Contrary to this, Suo et al. found no significant difference in D_K and K between grades 1, 2 and grade 3 [24].

4.3. Biases and limitations

The study was limited by the patient selection, as all included subjects had specific indications for an MRI scan (suspicious sonography or mammography, familiar risk, or history of breast cancer). Thus, the patient cohort cannot be considered equivalent, for example, to a screening population. Another limitation is that only a few benign lesions were analysed histologically (6/19 fibroadenomas, 2/41 cysts, 0/5 haematomas, 4/8 other benign lesions) as for ethical reasons, it was deemed unnecessary for the patients to undergo biopsy when no malignancy was suspected from the imaging results. However, follow-up scans after 12 months were consulted for classification of the lesions, which is an accepted approach in recent literature (e.g. [41]). Furthermore, the number of DCIS, Muc, and IntrCystPap lesions was small compared to the number of ILC and NST, and a larger patient cohort would be needed to verify the respective cut-off and differentiation performance values.

Importantly, the inclusion of cysts and haematoma as benign lesions can lead to an overestimation of the parameters' differentiation ability. Containing mostly free fluids, these lesions are easily identifiable on DWI and DKI sequences. We addressed this problem by separately evaluating the differentiation of malignant lesions from fibroadenomas, which are solid benign lesions. This provides a clinically more relevant differentiation performance. The resulting AUCs were lower in total than when all benign lesions were included, but the trend that ADC performed better than D_K and K remained.

Owing to the limited fit stability, D_K and K values were calculated from the mean signal intensities of the ROI. Measuring more b values could help to achieve more reliable D_K and K values (see Appendix). This would, however, increase the measurement time. The protocol used in this study was intentionally set up to take <5 min at maximum. This ensures that the protocol can be easily implemented into clinical routine.

For the calculation of the lesions' mean diffusion parameters, only a small ROI in one slice was used, excluding necroses and avoiding partial-volume effects. As the tumors can be heterogeneous, the chosen area is not necessarily representative of the whole lesion [42–44]. Segmenting the tumor in 3D could improve the meaningfulness of the median values [45].

5. Conclusions

We evaluated the diffusion metrics ADC, D_K and K in malignant and benign breast lesions in a short protocol with three b values $\leq 1500\,s/mm^2$ and TA < 5 min. All three metrics were well-suited for differentiation between malignant and benign lesions. Different tumor grades were significantly different in ADC and D_K, which may have an impact on therapy planning and monitoring. However, the more sophisticated DKI parameters brought no increase in diagnostic performance. Therefore, our data indicate that the commonly used ADC suffices as a diagnostic tool in a clinical context.

Declaration of competing interest

None.

Acknowledgements

We thank the Imaging Science Institute (Erlangen, Germany) for providing us with measurement time. The present work was performed in partial fulfillment of the requirements for obtaining the degree “Dr. rer. biol. hum.”.

Funding

This work was supported by the Deutsche Forschungsgemeinschaft, Germany [grant number LA 2804/6-1, project 326944748].

Research ethics and patient consent

This study was performed in accordance with the Declaration of Helsinki and received approval by the local institutional review board (Ethikkommission der Friedrich-Alexander-Universität Erlangen-Nürnberg, reference number 317_17 B). Written informed consent was obtained from the patients prior to examination.

Appendix A

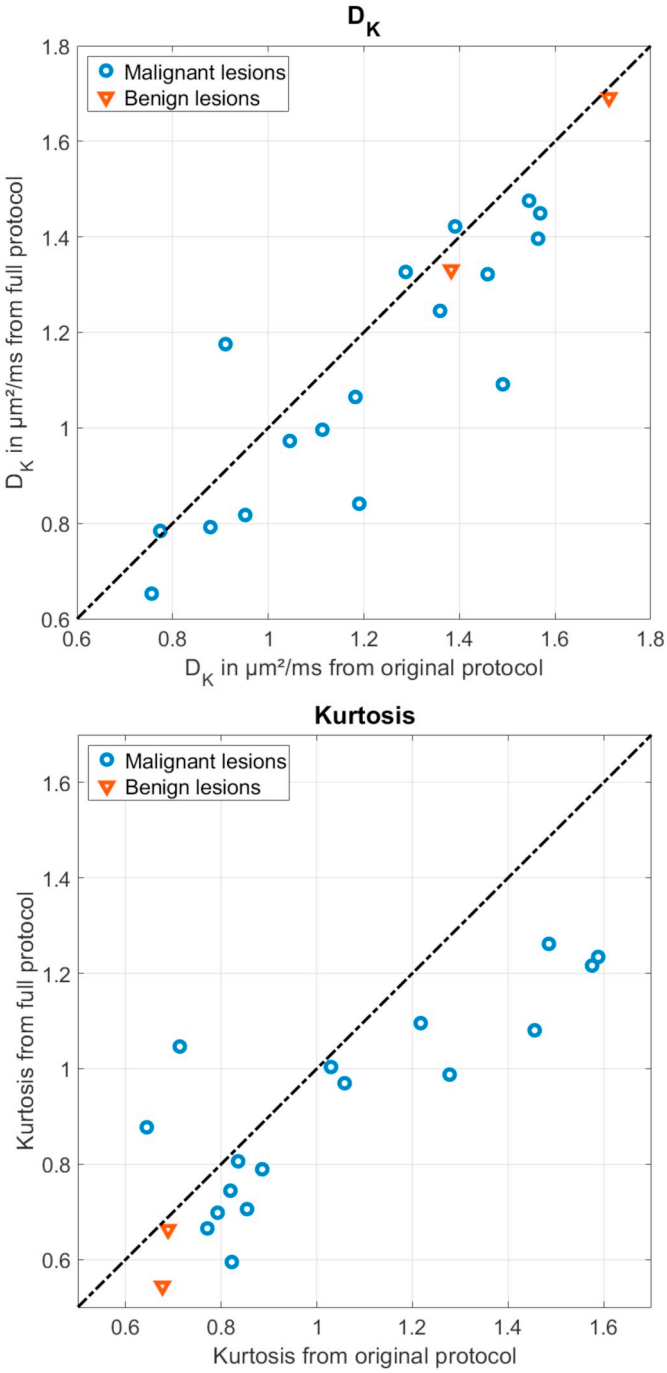

Fig. A1. Comparison of D_K and K values measured with the original and the full protocol. Malignant lesions are marked by blue circles, benign lesions by orange triangles. The dashed line indicates the bisector.

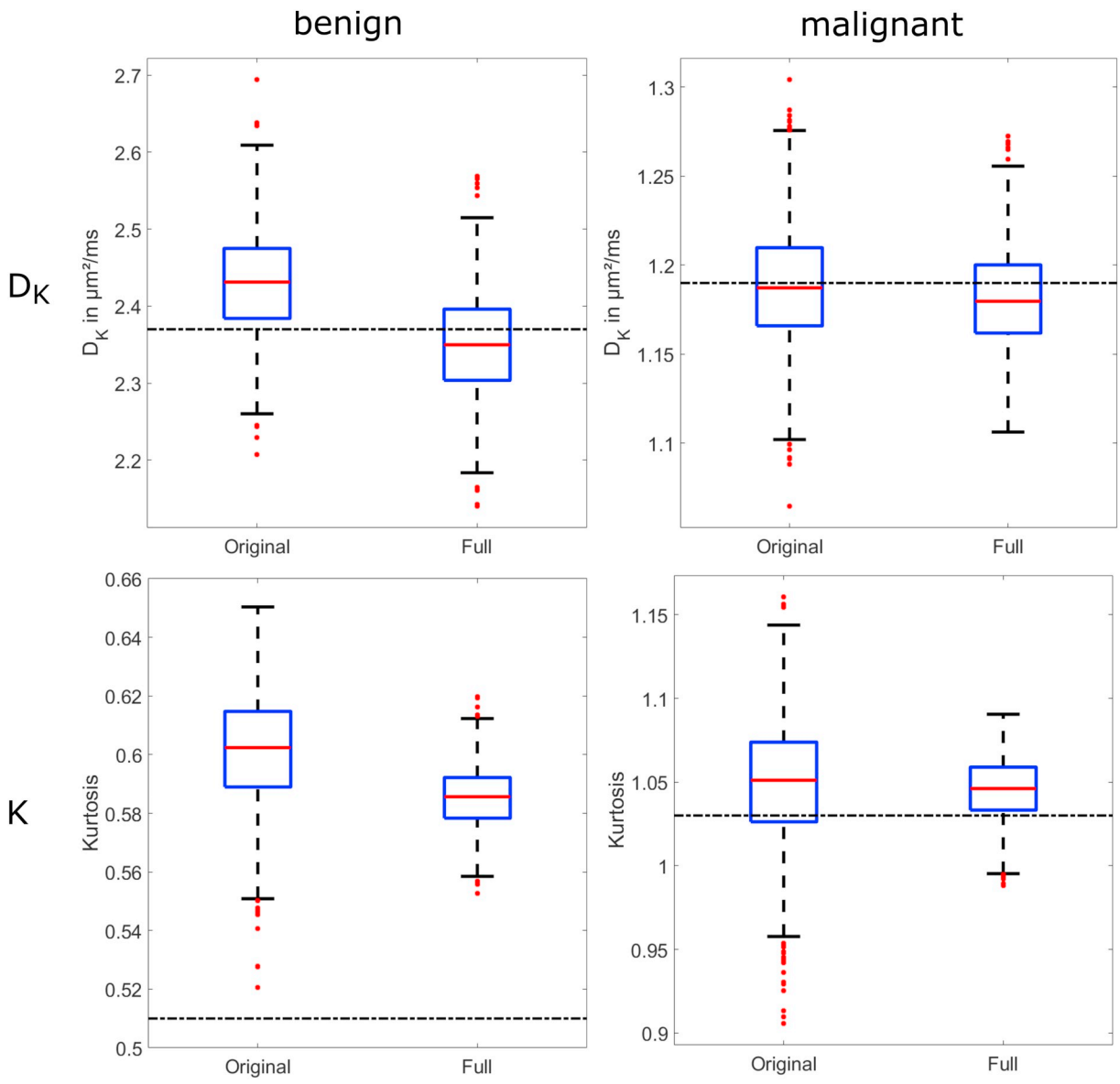

Fig. A2. Simulation fit results for original and full protocol. In 1000 lesions, signal values are simulated up to the second order of the cumulant expansion. Noise level 5 %, averaged over 14 pixels per ROI. Input parameters (indicated by dashed lines): benign: $D_{K,true}$ = 2.37 µm^2/ms, K_{true} = 0.51; malignant: $D_{K,true}$ = 1.19 µm^2/ms, K_{true} = 1.03.

Appendix B. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.mri.2019.08.007.

References

[1] Ghoncheh M, Pournamdar Z, Salehiniya H. Incidence and mortality and epidemiology of breast cancer in the world. Asian Pac J Cancer Prev 2016;17(S3):43–6.

[2] Saslow D, Boetes C, Burke W, Harms S, Leach Martin O, Lehman Constance D, et al. American Cancer Society guidelines for breast screening with MRI as an adjunct to mammography. CA Cancer J Clin 2009;57(2):75–89. https://doi.org/10.3322/canjclin.57.2.75.

[3] Lowry KP, Lee JM, Kong CY, McMahon PM, Gilmore ME, Cott Chubiz JE, et al. Annual screening strategies in BRCA1 and BRCA2 gene mutation carriers: a comparative effectiveness analysis. Cancer 2012;118(8):2021–30.

[4] Paluch-Shimon S, Pagani O, Partridge AH, Bar-Meir E, Fallowfield L, Fenlon D, et al. Second international consensus guidelines for breast cancer in young women (BCY2). Breast 2016;26:87–99.

[5] Orecchia R. MRI for treatment planning: a necessity. Eur J Radiol 2012;81:110–1.

[6] Pediconi F, Miglio E, Telesca M, Luciani ML, Kirchin MA, Passariello R, et al. Effect of preoperative breast magnetic resonance imaging on surgical decision making and cancer recurrence rates. Invest Radiol 2012;47(2):128–35. https://doi.org/10.1097/RLI.0b013e318230061c.

[7] Duygulu G, Oktay A, Bilgen IG, Kapkac M, Zekioglu O. The role of breast MRI in planning the surgical treatment of breast cancer. Diagn Interv Radiol 2012;18(5):460–7. https://doi.org/10.4261/1305-3825.Dir.5429-11.2.

[8] Mann RM, Kuhl CK, Kinkel K, Boetes C. Breast MRI: guidelines from the European Society of Breast Imaging. Eur Radiol 2008;18(7):1307–18. https://doi.org/10.1007/s00330-008-0863-7.

[9] Breast Imaging Working Group of the German Radiological Society. Updated recommendations for MRI of the breast. Fortschr Röntgenstr 2014;186(5):482–3. https://doi.org/10.1055/s-0034-1366404.

[10] D'Orsi CJ, Berg WA. ACR BI-RADS atlas: breast imaging reporting and data system. 4th ed American College of Radiology; 2013.

[11] McDonald RJ, McDonald JS, Kallmes DF, Jentoft ME, Murray DL, Thielen KR, et al. Intracranial gadolinium deposition after contrast-enhanced MR imaging. Radiology 2015;275(3):772–82. https://doi.org/10.1148/radiol.15150025.

[12] Kanda T, Ishii K, Kawaguchi H, Kitajima K, Takenaka D. High signal intensity in the dentate nucleus and globus pallidus on unenhanced T1-weighted MR images:

relationship with increasing cumulative dose of a gadolinium-based contrast material. Radiology 2013;270(3):834–41.
[13] Pullicino R, Radon M, Biswas S, Bhojak M, Das K. A review of the current evidence on gadolinium deposition in the brain. Clin Neuroradiol 2018;28:159–69.
[14] Saake M, Schmidle A, Kopp M, Hanspach J, Hepp T, Laun FB, et al. MRI brain signal intensity and relaxation times in individuals with prior exposure to gadobutrol. Radiology 2019;290(3):659–68. https://doi.org/10.1148/radiol.2018181927.
[15] Bickelhaupt S, Laun FB, Tesdorff J, Lederer W, Daniel H, Stieber A, et al. Fast and noninvasive characterization of suspicious lesions detected at breast cancer X-ray screening: capability of diffusion-weighted MR imaging with MIPs. Radiology 2016;278(3):689–97. https://doi.org/10.1148/radiol.2015150425.
[16] Bickelhaupt S, Paech D, Laun FB, Steudle F, Kuder TA, Mlynarska A, et al. Maximum intensity breast diffusion MRI for BI-RADS 4 lesions detected on X-ray mammography. Clin Radiol 2017;72(10):901–8. https://doi.org/10.1016/j.crad.2017.05.017.
[17] Guo Y, Cai YQ, Cai ZL, Gao YG, An NY, Ma L, et al. Differentiation of clinically benign and malignant breast lesions using diffusion-weighted imaging. J Magn Reson Imaging 2002;16(2):172–8. https://doi.org/10.1002/jmri.10140.
[18] Bogner W, Gruber S, Pinker K, Grabner G, Stadlbauer A, Weber M, et al. Diffusion-weighted MR for differentiation of breast lesions at 3.0 T: how does selection of diffusion protocols affect diagnosis? Radiology 2009;253(2):341–51. https://doi.org/10.1148/radiol.2532081718.
[19] EI Khouli RH, Jacobs MA, Mezban SD, Huang P, Kamel IR, Macura KJ, et al. Diffusion-weighted imaging improves the diagnostic accuracy of conventional 3.0-T breast MR imaging. Radiology 2010;256(1):64–73.
[20] Suo S, Zhang K, Cao M, Suo X, Hua J, Geng X, et al. Characterization of breast masses as benign or malignant at 3.0 T MRI with whole-lesion histogram analysis of the apparent diffusion coefficient. J Magn Reson Imaging 2016;43(4):894–902.
[21] Wenkel E, Geppert C, Schulz-Wendtland R, Uder M, Kiefer B, Bautz W, et al. Diffusion weighted imaging in breast MRI: comparison of two different pulse sequences. Acad Radiol 2007;14(9):1077–83.
[22] Iima M, Yano K, Kataoka M, Umehana M, Murata K, Kanao S, et al. Quantitative non-Gaussian diffusion and intravoxel incoherent motion magnetic resonance imaging: differentiation of malignant and benign breast lesions. Invest Radiol 2015;50(4):205–11.
[23] Jensen JH, Helpern JA, 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 2005;53(6):1432–40. https://doi.org/10.1002/mrm.20508.
[24] Suo S, Cheng F, Cao M, Kang J, Wang M, Hua J, et al. Multiparametric diffusion-weighted imaging in breast lesions: association with pathologic diagnosis and prognostic factors. J Magn Reson Imaging 2017;46(3):740–50. https://doi.org/10.1002/jmri.25612.
[25] Iima M, Kataoka M, Kanao S, Onishi N, Kawai M, Ohashi A, et al. Intravoxel incoherent motion and quantitative non-Gaussian diffusion MR imaging: evaluation of the diagnostic and prognostic value of several markers of malignant and benign breast lesions. Radiology 2017;287(2):432–41. https://doi.org/10.1148/radiol.2017162853.
[26] Wu D, Li G, Zhang J, Chang S, Hu J, Dai Y. Characterization of breast tumors using diffusion kurtosis imaging (DKI). PLoS One 2014;9(11):e113240https://doi.org/10.1371/journal.pone.0113240.
[27] Sun K, Chen X, Chai W, Fei X, Fu C, Yan X, et al. Breast cancer: diffusion kurtosis MR imaging—diagnostic accuracy and correlation with clinical-pathologic factors. Radiology 2015;277(1):46–55.
[28] Nogueira L, Brandao S, Matos E, Nunes RG, Loureiro J, Ramos I, et al. Application of the diffusion kurtosis model for the study of breast lesions. Eur Radiol 2014;24(6):1197–203. https://doi.org/10.1007/s00330-014-3146-5.
[29] Christou A, Ghiatas A, Priovolos D, Veliou K, Bougias H. Accuracy of diffusion kurtosis imaging in characterization of breast lesions. Br J Radiol 2017;90(1073):20160873https://doi.org/10.1259/bjr.20160873.
[30] Huang Y, Lin Y, Hu W, Ma CC, Lin WX, Wang ZN, et al. Diffusion kurtosis at 3.0 T as an in vivo imaging marker for breast cancer characterization: correlation with prognostic factors. J Magn Reson Imaging 2019;49(3):845–56. https://doi.org/10.1002/jmri.26249.
[31] Li T, Yu T, Li L, Lu LB, Zhuo YY, Lian JG, et al. Use of diffusion kurtosis imaging and quantitative dynamic contrast-enhanced MRI for the differentiation of breast tumors. J Magn Reson Imaging 2018;48(5):1358–66. https://doi.org/10.1002/jmri.26059.
[32] Wallis M, Tarvidon A, Helbich T, Schreer I. Guidelines from the European Society of Breast Imaging for diagnostic interventional breast procedures. Eur Radiol 2007;17(2):581–8.
[33] R Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing; 2018.
[34] DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 1988;44(3):837–45.
[35] Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez J-C, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinf 2011;12(1):77. https://doi.org/10.1186/1471-2105-12-77.
[36] Woodhams R, Kakita S, Hata H, Iwabuchi K, Umeoka S, Mountford CE, et al. Diffusion-weighted imaging of mucinous carcinoma of the breast: evaluation of apparent diffusion coefficient and signal intensity in correlation with histologic findings. AJR Am J Roentgenol 2009;193(1):260–6. https://doi.org/10.2214/ajr.08.1670.
[37] Woodhams R, Matsunaga K, Kan S, Hata H, Ozaki M, Iwabuchi K, et al. ADC mapping of benign and malignant breast tumors. Magn Reson Med Sci 2005;4(1):35–42.
[38] Janka R, Wenkel E, Geppert C. Diffusion weighted imaging of the breast: comparison of different pulse sequences in breast lesions at 3.0 T. Magma ESMRMB. 412. 2009.
[39] Iima M, Kataoka M, Kanao S, Kawai M, Onishi N, Koyasu S, et al. Variability of non-Gaussian diffusion MRI and intravoxel incoherent motion (IVIM) measurements in the breast. PLoS One 2018;13(3):12. https://doi.org/10.1371/journal.pone.0193444.
[40] Rahbar H, Partridge SC, Eby PR, DeMartini WB, Gutierrez RL, Peacock S, et al. Characterization of ductal carcinoma in situ on diffusion weighted breast MRI. Eur Radiol 2011;21(9):2011–9. https://doi.org/10.1007/s00330-011-2140-4.
[41] Förnvik D, Kataoka M, Iima M, Ohashi A, Kanao S, Toi M, et al. The role of breast tomosynthesis in a predominantly dense breast population at a tertiary breast centre: breast density assessment and diagnostic performance in comparison with MRI. Eur Radiol 2018:1–10.
[42] Partridge SC, Amornsiripanitch N. DWI in the assessment of breast lesions. Top Magn Reson Imaging 2017;26(5):201–9. https://doi.org/10.1097/rmr.0000000000000137.
[43] Bickel H, Pinker K, Polanec S, Magometschnigg H, Wengert G, Spick C, et al. Diffusion-weighted imaging of breast lesions: region-of-interest placement and different ADC parameters influence apparent diffusion coefficient values. Eur Radiol 2017;27(5):1883–92.
[44] Hering J, Laun FB, Lederer W, Daniel H, Kuder TA, Stieber A, et al. Applicability and discriminative value of a semiautomatic three-dimensional spherical volume for the assessment of the apparent diffusion coefficient in suspicious breast lesions-feasibility study. Clin Imaging 2016;40(6):1280–5. https://doi.org/10.1016/j.clinimag.2016.08.022.
[45] Rahbar H, Kurland BF, Olson ML, Kitsch AE, Scheel JR, Chai X, et al. Diffusion-weighted breast magnetic resonance imaging: a semiautomated voxel selection technique improves interreader reproducibility of apparent diffusion coefficient measurements. J Comput Assist Tomogr 2016;40(3):428–35.