Eur Radiol (2017) 27:4336–4344
DOI 10.1007/s00330-017-4814-z

UROGENITAL

Selection for biopsy of kidney transplant patients by diffusion-weighted MRI

Philipp Steiger[1] · **Sebastiano Barbieri**[1] · **Anja Kruse**[2] · **Michael Ith**[1] · **Harriet C. Thoeny**[1]

Received: 10 September 2016 / Revised: 26 February 2017 / Accepted: 14 March 2017 / Published online: 3 April 2017

Abstract

Objectives To assess retrospectively whether diffusion-weighted magnetic resonance imaging (DW-MRI) allows physicians to determine the severity of histopathologic findings in biopsies of renal allograft patients with deteriorating renal function.

Methods Forty consecutive kidney transplant patients underwent DW-MRI and biopsy. Patients were assigned to one group with severe and to another group with normal or mild histopathologic findings. These two groups were compared based on a *qualitative* DW-MRI assessment (homo-/heterogeneity) and the combination of *qualitative* and *quantitative* DW-MRI parameters (ADC, and intravoxel incoherent motion, IVIM, parameters: D, f, D*). Sensitivity, specificity, and accuracy were determined for each parameter.

Results Biopsy findings were severe in 25 patients and normal or mild in 15 patients. Q*ualitative* DW-MRI led to a sensitivity of 44.0% and a specificity of 93.3%. Combined *qualitative* and *quantitative* DW-MRI led to an accuracy of 80% for both the minimal ADC (ADCmin) and the minimal perfusion fraction (*f*min) with a sensitivity of 84.0% and 92.0% and a specificity of 73.3% and 60.0%, respectively.

Conclusion Combined *qualitative* and *quantitative* DW-MRI might allow physicians to determine the severity of histopathologic findings in biopsies of a high number of kidney transplant patients.

Key points

- *Qualitative DW-MRI is highly specific when predicting the severity of kidney transplant biopsy.*
- *Allografts appearing heterogeneous on ADC are associated with severe histopathologic findings.*
- *Combining qualitative and quantitative DW-MRI parameters improves the classification's sensitivity and accuracy.*
- *Kidney transplant biopsies might be spared by combining qualitative and quantitative DW-MRI.*

Keywords Renal transplant · Diffusion-weighted magnetic resonance imaging · Echo-planar imaging · Biopsy · Anatomy & histology

Abbreviations

DW-MRI	Diffusion-weighted MRI
Mild histopathologic changes	Grade I interstitial fibrosis and tubular atrophy
Severe histopathologic changes	Grades II-III interstitial fibrosis and tubular atrophy rejection, and other changes not due to rejection such as glomerulonephritis, IgA nephropathy, BK virus nephropathy, acute tubular necrosis, and cyclosporine toxicity
ADC	Apparent diffusion coefficient

Electronic supplementary material The online version of this article (doi:10.1007/s00330-017-4814-z) contains supplementary material, which is available to authorized users.

✉ Harriet C. Thoeny
Harriet.thoeny@insel.ch

[1] Department of Radiology, Neuroradiology, and Nuclear Medicine, Institute of Diagnostic, Pediatric, and Interventional Radiology, Inselspital, Bern University Hospital, University of Bern, Inselspital, Freiburgstrasse 10, CH-3010 Bern, Switzerland

[2] Department of Nephrology and Hypertension, Inselspital, Bern University Hospital, University of Bern, Inselspital, Freiburgstrasse 10, CH-3010 Bern, Switzerland

IVIM	Intravoxel incoherent motion
f	Perfusion fraction
*D**	Pseudo-diffusion coefficient
D	True diffusion coefficient
s-Crea	Serum creatinine
ROIs	Regions of interest
ROC	Receiver operating characteristic

Introduction

Kidney allograft recipients are currently treated, managed, and monitored according to their overall condition, as well as their blood and urine laboratory tests. Biopsies of kidney allografts might be performed when one of the following conditions occurs: unexplained increase in serum creatinine (s-Crea); s-Crea has not returned to baseline after treatment of acute rejection; delayed graft function; expected graft function is not achieved within 1–2 months after transplantation; new onset of proteinuria or unexplained proteinuria. Although kidney transplant biopsies are a frequently performed and generally safe procedure, complications still occur. In protocol biopsies, the reported risk of major complications, including substantial bleeding, macroscopic haematuria with ureteric obstruction, and peritonitis, is approximately 1%. The prevalence of graft loss is reported at 0.03% [1–3], and it is currently unclear whether the benefits of protocol and control biopsies outweigh their harm [4].

Kidney transplant biopsies are evaluated according to the Banff classification [5]. At our institution, a biopsy result showing normal or mild histopathologic findings had generally no direct consequence on the patient's clinical management. On the contrary, higher grade (II-III) interstitial fibrosis and tubular atrophy, rejection and other changes not due to rejection (glomerulonephritis, IgA nephropathy, BK virus nephropathy, acute tubular necrosis and cyclosporine toxicity) often had direct implications on the patient's clinical management. Specifically, the detection of interstitial fibrosis in the transplant may lead, depending on severity and suspected cause (e.g. calcineurin inhibitor toxicity or chronic humoral rejection), to an adaptation of the immunosuppressive therapy.

Colour Doppler ultrasound is the most used imaging procedure in kidney transplant patients, but has several limitations such as interobserver variability, imaging difficulties in obese patients, and limited resolution in diffuse parenchymal pathologies [6]. Although magnetic resonance imaging (MRI) of transplanted kidneys is not yet part of daily clinical routine, diffusion-weighted MRI (DW-MRI) without contrast medium administration has shown promising results in the detection of microstructural and functional changes of renal allografts before morphological differences become evident [7–10]. This noninvasive technique is gaining importance especially for patients who have renal impairment and for whom the administration of MR contrast agents has become a matter of concern in recent years: these patients are at higher risk of developing nephrogenic systemic fibrosis [11].

In transplanted kidneys, DW-MRI is a feasible and reproducible technique which allows to analyse quantitatively the displacement of water molecules in the underlying tissue by computing the apparent diffusion coefficient (ADC) [6]. Further, the intravoxel incoherent motion (IVIM) model allows physicians to measure the perfusion fraction (*f*), the pseudo-diffusion coefficient (*D**), and the true diffusion coefficient (*D)* [12]. Several studies have reported that microstructural changes occurring in histopathologic alterations correlate with changes in ADC, *f,* and *D* [6–8].

The purpose of this study is to assess retrospectively whether the severity of histopathologic findings in kidney allograft biopsies (based on implications on the patient's clinical management) can be determined by combining *qualitative* (kidney homo-/heterogeneity) and *quantitative* (ADC, *f*, *D**, and *D*) DW-MRI parameters.

Materials and methods

Patients, histopathologic findings, and their classification

The local ethics committee waived the approval requirement for this retrospective study. A prospective study evaluating the applied protocol has been performed and published previously [7]. All patients sent for MRI of renal transplants due to deteriorating renal function were examined according to this protocol. Although the MRI protocol was initially applied prospectively, image analysis of those patients undergoing renal biopsy has been performed retrospectively in the present study. The study included 40 kidney transplant patients (30 men, 10 women; mean age at biopsy ± standard deviation (SD), 58.5 ± 13.0 years) with deteriorating renal function. These patients underwent biopsy of their renal allografts within 10 days before or after MR imaging. Patient characteristics and histopathologic findings are summarized in Table 1. Patients were divided into two groups: the "normal or mild histopathologic changes" group included patients whose histopathologic findings did not have any consequence on the patients' clinical management (normal or mild); the "severe histopathologic changes" group included patients whose histopathologic findings had implications on the patients' clinical management (grades II-III interstitial fibrosis and tubular atrophy, rejection, and other changes not due to rejection such as glomerulonephritis, IgA nephropathy, BK virus nephropathy, acute tubular necrosis, and cyclosporine toxicity).

MR imaging

MR imaging was performed on a 1.5-T MR unit (Sonata; Siemens, Erlangen, Germany) using a six-channel body coil.

Table 1 Patient demographics

Age at MR imaging (years) †	55 (24 – 70)
Time between transplantation and MR (days) †	679 (1 – 6799)
Time between biopsy and MR (days) †	1 (-7 – 10)
Mean s-Crea at time of biopsy or MR (μmol/L) †	258 (85 – 832)
Mean s-Crea at time of biopsy or MR in different histopathologic findings subgroup (μmol/L) †:	
Normal or mild histopathologic findings:	225 (85 – 504)
Severe histopathologic findings:	271 (138 – 832)
Histologic diagnosis according to the Banff classification	
Normal or mild histopathologic findings:	
Normal	1
Interstitial fibrosis and tubular atrophy grade I (exclusively grade I) §	30 (14)
Severe histopathologic findings:	
Antibody mediated rejection	2
Borderline rejection	2
T-cell mediated rejection	6
Interstitial fibrosis and tubular atrophy (grade II/III)	8 (8/0)
Other, consisting of:	18
Glomerulonephritis	1
IgA nephropathy	2
BK virus nephropathy	2
Acute tubular necrosis	3
Cyclosporine toxicity	10

† Values are medians with range in parenthesis. § Sixteen patients with normal and grade I mild interstitial fibrosis and tubular atrophy had additional severe histopathology findings. Note: Some patients had more than one histologic diagnosis; for details see supplemental material.

For morphological evaluation, transverse T2-weighted (time of repetition, TR, 1100 ms; echo time, TE, 114 ms) and transverse and coronal T1-weighted (TR 4.76 ms, TE 74 ms) images were acquired. For functional evaluation, coronal multisection echo-planar DW-MRI was performed with the following parameters: 12-15 sections (section thickness, 5 mm; intersection gap, 1 mm); field of view, 400 × 400 mm; matrix, 128 × 128; TR, 3200 ms; TE, 71 ms; six signals acquired; bandwidth, 1500 Hz per pixel; and partial Fourier factor 6/8. The following 10 diffusion gradient b-values were used: 0, 10, 20, 40, 60, 150, 300, 500, 700, and 900 s/mm^2. The gradients where applied in three orthogonal directions and subsequently averaged to reduce the effects of diffusion anisotropy. A parallel imaging technique (modified sensitivity encoding) with a reduction factor of two was applied. Respiratory triggering was used to reduce motion artefacts. Slice positioning was identical to the one used for the coronal T1-weighted sequence. The minimum acquisition time for DW-MRI was 9 min 20 s.

Image analysis

Both the morphological T1- and T2-weighted images and the functional DW-MR images were evaluated in consensus by two radiologists (H.C.T., 20 and 13 years of experience with renal MRI and DW-MRI, respectively; P.S., 1 year of experience with DW-MRI) who were blinded to the histopathologic results. Conventional MR images were visually analysed to detect any morphologic abnormalities including reduced/absent corticomedullary differentiation and focal lesions.

The DW-MR images were analysed *qualitatively* and then *quantitatively*.

The *qualitative* image analysis evaluated whether the whole kidney appeared homogeneous or heterogeneous (i.e. whether focal changes were visible or not) on high b-value images (b = 900 s/mm^2) and on the corresponding ADC maps (Fig. 1).

Quantitative parameters were derived from DW-MR images by means of a custom ImageJ plugin as previously described [7, 13] on a pixel-by-pixel basis using all acquired b-values. A monoexponential fit was used to compute the ADC values and a biexponential fit (using the Levenberg-Marquardt algorithm) was used to determine the values of the IVIM parameters (D, f, and D^*). For every patient with a homogeneous appearing allograft, ellipsoid regions of interest (ROIs) of four to six pixels each were placed on a coronal slice of the ADC map of the transplanted kidney. A total of three ROIs were placed: one in the upper, one in the middle, and one in the lower pole of the kidney. Parameters were averaged across pixels and subsequently across ROIs. The resulting values are denoted as ADCmean, *D*mean, *D**mean, and *f*-mean. The lowest parameter values across the three ROIs are denoted as ADCmin, *D*min, *D**min, and *f*min.

In patients with heterogeneous appearing allografts only qualitative image analysis has been performed.

Cortex and medulla were not considered separately for ROI positioning because they were indistinguishable on the ADC maps. Care was taken not to include cystic lesions or the renal pelvis within the ROIs.

Statistical analysis

In homogeneous appearing kidneys two-sample t-tests were used to analyse whether ADC and IVIM parameter values differed between “severe histopathologic changes” and “normal or mild histopathologic changes”.

Patients were classified as having allografts with “severe” or “normal and mild histopathologic changes” based on kidney homo-/heterogeneity alone and on kidney homo-/heterogeneity combined with ADC and IVIM parameter values which were used to refine the classification of patients with homogeneous appearing kidneys (Fig. 2). Receiver operating characteristic (ROC) curves were used to assess the sensitivity, specificity, accuracy, and area under the curve (AUC) of these binary classifiers.

Two different threshold values for ADC and IVIM parameter values were determined: the first set of threshold values, denoted as t-MaxSum, maximizes the sum of sensitivity and

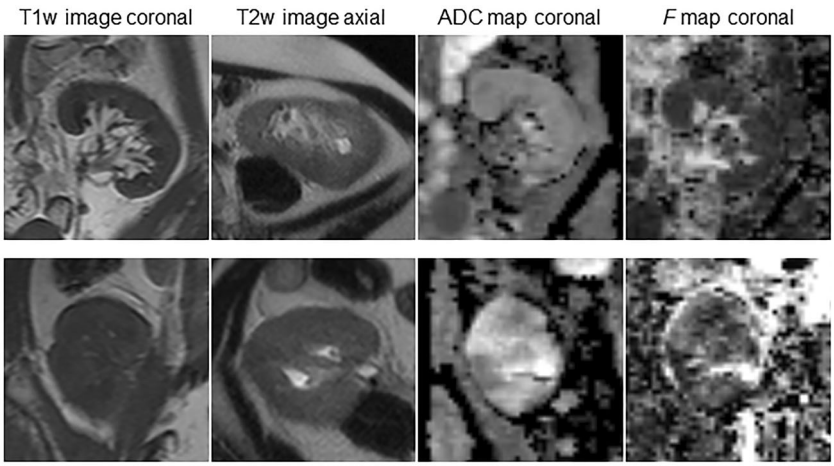

Fig. 1 *Upper row:* Patient in the "normal or mild histopathologic changes" group with mild histopathologic changes showing a homogenous ADC and *f*map. *Lower row:* Patient in the "severe histopathologic changes" group with an acute tubular necrosis and a BK virus nephropathy showing a heterogeneous ADC and *f* map. On morphologic T1w and T2w images the kidneys of patients in the two groups can not be distinguished

specificity of the classifiers; the second set of threshold values, denoted as t-MaxSens, maximizes the specificity of the classifiers when their sensitivity is 100%. The latter threshold values are of interest because of the high cost associated with not performing a biopsy that would have had implications on the patient's clinical management.

Creatinine levels between the "severe histopathologic changes" group and the "normal or mild histopathologic changes" group were compared by a two-sample t-test.

For all statistical tests, a P value of less than .05 was assumed to indicate statistical significance.

Statistical analyses were performed with SPSS, version 12.0 (SPSS, Chicago, IL, USA), and Excel 2010 (Microsoft, Redmond, WA, USA) software.

Results

Of the 40 patients, 15 were in the "normal or mild histopathologic changes" group, where kidney transplant biopsy revealed normal (1) or merely mild histopathologic changes (14). Twenty-five patients were in the "severe histopathologic changes" group, where histopathologic findings revealed higher grade (II-III) interstitial fibrosis and tubular atrophy (eight grade II and zero grade III), rejection [antibody mediated rejection (2), borderline rejection (2), T-cell mediated rejection (6)] and other changes not due to rejection [glomerulonephritis (1), IgA nephropathy (2), BK virus nephropathy (2), acute tubular necrosis (3), and cyclosporine toxicity (10)]. Kidneys in the "severe histopathologic changes" group were often associated with several different findings on histopathologic analysis, details are reported in Table 1.

Morphology on T1- and T2-weighted sequences showed no differences between the "severe histopathologic changes" and the "normal or mild histopathologic changes" group. Further, none of the subcategories of the Banff classification system were associated with specific morphological image patterns.

An internal quality control of the *qualitative* image analysis step indicated, as expected, that the signal intensity variance within ROIs placed in heterogeneous kidneys was

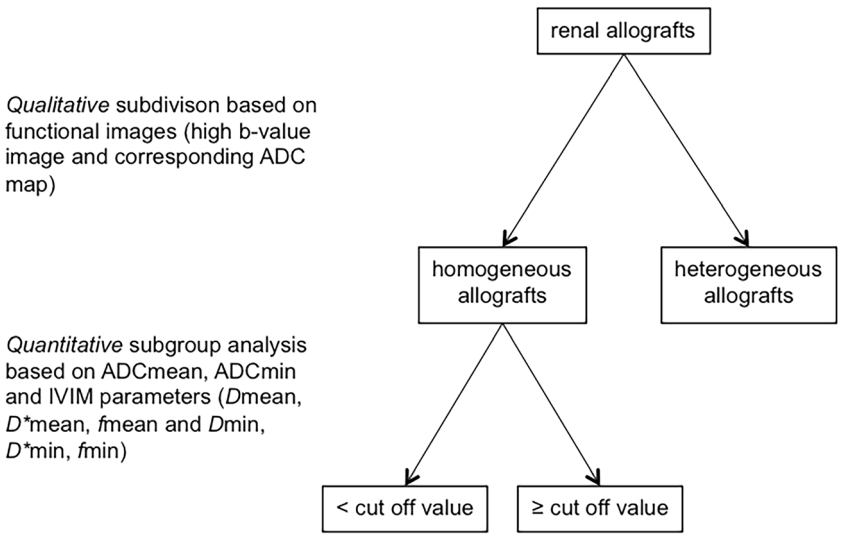

Fig. 2 Algorithm for the classification of imaged allografts based on a combined *qualitative* and *quantitative* DW-MRI analysis. In a first step kidneys are subdivided into homo- or heterogeneous depending on their appearance on high b-value images and on the corresponding ADC map. Next, homogenous kidneys are separated *quantitatively* based on the ADC-value or based on IVIM parameter values (*D*mean, *D**mean, *f*mean and *D*min, *D**min, *f*min)

considerably higher than in ROIs placed in homogeneous kidneys (see supplemental material).

In homogeneous kidneys, the two-sample t-tests did not indicate that ADC or IVIM parameter values differed significantly between the "severe histopathologic changes" group and the "normal or mild histopathologic changes" group; however, some evidence suggested a difference in *f*min ($P = 0.04$, see Table 2). Assigning kidneys to the "severe histopathologic changes" or the "normal or mild histopathologic changes" group based solely on their homo-/heterogeneity resulted in a sensitivity of 44.0% and a specificity of 93.3%, with an accuracy of 62.5%.

Quantitative image information (ADC and IVIM parameter values measured in homogenous appearing kidneys) in combination with *qualitative* information (homo-/heterogeneity of kidneys) yielded the highest accuracy (80.0%) when homogeneous kidneys were classified according to their ADCmin and *f*min values (sensitivity 84.0% and 92.0%; specificity 73.3% and 60.0%, respectively, using the t-MaxSum threshold criterion, see Table 3). Boxplots and corresponding ROC curves for ADCmin and *f*min values are shown in Fig. 3. Maximisation of sensitivity to 100% was reached by increasing threshold values for ADC and IVIM parameters to the point where no "severe histopathologic changes" patient would have been missed for biopsy (t-MaxSens threshold criterion). Using the t-MaxSens criterion the highest specificity and accuracy values were determined for the *f*mean and *f*min parameters (specificity of 40.0% and accuracy of 77.5% for both, see Table 4). Therefore, classification based on *f*mean or *f*min parameter values would have spared unnecessary biopsies to six out of 15 patients. These six patients were the same when using either *f*mean or *f*min and the two negative patients who were classified correctly by ADCmean and ADCmin were among these six patients as well.

A two-sample t-test showed no statistically significant difference in creatinine levels between the "severe histopathologic changes" group and the "normal or mild histopathologic changes" group ($P = 0.30$).

Discussion

The present study shows that DW-MRI allows predicting the severity of biopsies in renal transplant patients with deteriorating renal function. Specifically, patients might be selected for biopsy based on the combined analysis of *qualitative* (kidney homo-/heterogeneity) and *quantitative* (ADC, *f*, *D**, and *D* measured in the homogenous subgroup) DW-MR image analysis. This is an improvement over other non-invasive diagnostic approaches that have been suggested to monitor graft dysfunction (e.g. ultrasound imaging [14]), but where the observed characteristic occurs late or is non-specific.

In accordance with a previously published study [10], kidney transplants with a heterogeneous appearance on ADC maps were associated with severe histopathologic changes. The high specificity of transplant classification based on heterogeneous appearance might be explained by the way in which pathological conditions affect kidney transplants: from the histopathologic analysis of native kidneys it is known that diseases can affect kidneys focally or in their entirety [15]; the former possibility might, therefore, result in a heterogeneous image pattern of the kidney transplant on DW-MRI. Consequently, we suggest that every heterogeneous appearing kidney transplant on DW-MRI should undergo biopsy.

However, the sensitivity of classification based on heterogeneous appearance is insufficient for heterogeneity to be used as the sole criterion to determine any severe histopathologic changes in kidney transplant biopsies. Nonetheless,

Table 2 Mean difference of ADC and IVIM parameter values in homogeneous kidneys

	Mean difference ("normal or mild histopathologic changes" – "severe histopathologic changes")	95% Confidence interval		*P* of t-test
		Lower	Upper	
ADCmean [10^{-5} mm^2/s]	11.2	-4.3	26.7	0.15
*D**mean [10^{-4} mm^2/s]	-30	-73.5	21.1	0.167
*D*mean [10^{-6} mm^2/s]	72.4	-58.2	202.9	0.265
*f*mean [%]	4.3	-0.2	8.8	0.061
ADCmin [10^{-5} mm^2/s]	12.6	-2.5	27.6	0.098
*D**min [10^{-4} mm^2/s]	-4.8	-31.9	21.7	0.072
*D*min [10^{-6} mm^2/s]	58.6	-69	186.1	0.354
*f*min [%]	4.4	0.2	8.6	0.042

Mean difference of ADC and IVIM parameter values between patients in the "normal or mild histopathologic changes" and the "severe histopathologic changes" group. Patients in the "severe histopathologic changes" group showed lower ADC and IVIM parameter values except for D*. A two-sample t-test determined a statistically significant difference only for *f*min.

Table 3 ROC analysis for different DW-MRI criteria when predicting severe histopathologic changes in renal allograft biopsy

DW-MRI criteria for severe histopathologic changes	Sensitivity (%)	Specificity (%)	Positive predictive value (%)	Negative predictive value (%)	Accuracy (%)	Corresponding ROC cutoff and AUC	
						Cutoff value	ROC AUC
A) Heterogeneous	44.0 (11/25)	93.3 (14/15)	91.7 (11/12)	50.0 (14/28)	62.5 (25/40)		
B) Homogeneous with a cutoff value for							
ADCmean [10^{-5} mm^2/s]	71.4 (10/14)	71.4 (10/14)	71.4 (10/14)	71.4 (10/14)	71.4 (20/28)	188.8	0.668
D*mean [10^{-4} mm^2/s]	64.3 (9/14)	71.4 (10/14)	69.2 (9/13)	66.7 (10/15)	67.9 (19/28)	332.5	0.694
*D*mean [10^{-6} mm^2/s]	71.4 (10/14)	57.1 (8/14)	62.5 (10/16)	66.7 (8/12)	64.3 (18/28)	1850	0.597
*f*mean [%]	64.3 (9/14)	64.3 (9/14)	64.3 (9/14)	64.3 (9/14)	64.3 (18/28)	16.3	0.709
ADCmin [10^{-5} mm^2/s]	71.4 (10/14)	78.6 (11/14)	76.9 (10/13)	73.3 (11/15)	75.0 (21/28)	179.5	0.691
*D**min [10^{-4} mm^2/s]	57.1 (8/14)	78.6 (11/14)	72.7 (8/11)	64.7 (11/17)	67.9 (19/28)	266.5	0.602
*D*min [10^{-6} mm^2/s]	64.3 (9/14)	64.3 (9/14)	64.3 (9/14)	64.3 (9/14)	64.3 (18/28)	1735	0.61
*f*min [%]	85.7 (12/14)	64.3 (9/14)	70.6 (12/17)	81.8 (9/11)	75.0 (21/28)	14	0.719
Combination of heterogeneous and homogeneous (A&B)							
ADCmean [10^{-5} mm^2/s]	84.0 (21/25)	66.7 (10/15)	80.8 (21/26)	71.4 (10/14)	77.5 (31/40)	188.8	0.668
*D**mean [10^{-4} mm^2/s]	80.0 (20/25)	66.7 (10/15)	80.0 (20/25)	66.7 (10/15)	75.0 (30/40)	332.5	0.694
*D*mean [10^{-6} mm^2/s]	84.0 (21/25)	53.0 (8/15)	75.0 (21/28)	66.7 (8/12)	72.5 (29/40)	1850	0.597
*f*mean [%]	80.0 (20/25)	60.0 (9/15)	76.9 (20/26)	64.3 (9/14)	72.5 (29/40)	16.3	0.709
ADCmin [10^{-5} mm^2/s]	84.0 (21/25)	73.3 (11/15)	84.0 (21/25)	73.3 (11/15)	80.0 (32/40)	179.5	0.691
*D**min [10^{-4} mm^2/s]	76.0 (19/25)	73.3 (11/15)	82.6 (19/23)	64.7 (11/17)	75.0 (30/40)	266.5	0.602
*D*min [10^{-6} mm^2/s]	80.0 (20/25)	60.0 (9/15)	76.9 (20/26)	64.3 (9/14)	72.5 (29/40)	1735	0.61
*f*min [%]	92.0 (23/25)	60.0 (9/15)	79.3 (23/29)	81.8 (9/11)	80.0 (32/40)	14	0.719

antibody-mediated rejection, borderline rejection, and BK virus nephropathy were detected exclusively in heterogeneous appearing kidneys.

Including *quantitative* DW-MRI parameters in our analysis increased the sensitivity and accuracy for ADC and all IVIM parameters (highest accuracy for ADCmin and *f*min). Minimal ROI values discriminate "severe histopathologic changes" and "normal or mild histopathologic changes" patients better than mean ROI values. An explanation could be that in "severe histopathologic changes" patients, minor pathologic changes result in diffusion and perfusion measurements that are altered locally, but not on average for the whole transplant. In the model with maximisation of sensitivity, the perfusion fraction *f* was the best discriminator between "severe histopathologic changes" and "normal or mild histopathologic changes" patients. These findings are in line with a previous study in a rat model [16], which showed that *f* is the most sensitive IVIM parameter for the detection of renal pathologies, explained by a decrease of *f* in pathologic kidneys due to both the reduced capillary vasculature and the narrowed renal tubules. In addition, as has been shown by kidney transplant angiograms, perfusion can be altered because of organ rejection [17]. This study also shows that rejection can be a predominantly focal process, with regions of the kidneys that are less or even completely unaffected. The described results are also consistent with our own experience that kidneys with heterogeneous appearance on DW-MRI are affected by an underlying pathology. A biopsy of a kidney transplant allows examining only a small specimen of the transplant kidney, whereas on DW-MRI the whole organ can be investigated. Therefore, it might be speculated that kidney transplant biopsy alone is not the optimal evaluation tool for the surveillance of kidney transplant patients. DW-MRI might serve to monitor renal integrity (together with the course of s-Crea) and as a guide for taking biopsy samples in case of focal change.

We are aware of only one study that assessed the correlation between ADC maps of kidneys with impaired function and histologic findings [10]. It was shown that ADC is lower in transplanted kidneys with impaired function compared to kidneys with stable function. These results cannot be directly compared with ours as all our patients had unstable renal function and underwent biopsy. However, lower ADC values and a characteristic heterogeneous mosaic pattern on ADC maps were found to be associated with acute tubular necrosis [10], while in our study the kidney allograft of one of the three patients with acute tubular necrosis appeared homogeneous on DW-MRI. Additionally, our study did not indicate that heterogeneous appearance on DW-MRI was associated with a specific pathology of the transplant.

The number of kidney transplant biopsies is reduced in a non-control biopsy setting (as in our study) compared with a control biopsy setting, which increases the risk of missing

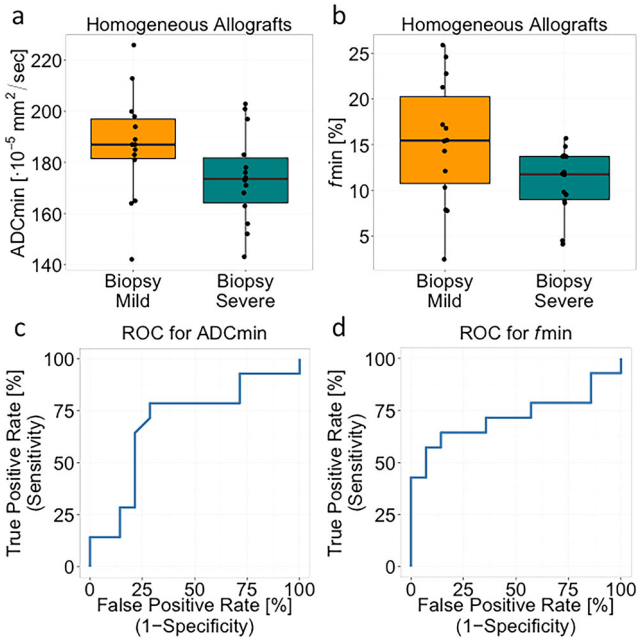

Fig. 3 Transplants with homogeneous appearance on ADC maps can be classified as "normal or mild histopathologic changes" or "severe histopathologic changes" by a cutoff value for ADCmin of 179.5×10^{-5} mm^2/s (**a**) or for *f*min of 14.0% (**b**). Corresponding ROC curves for ADCmin and *f*min are depicted in (**c**) and (**d**)

subclinical histological changes [18]. It is not known whether these subclinical changes might affect the patient's therapy. The maximisation of the classifier's sensitivity to 100% indicated that cutoff values for *f*mean and *f*min in homogeneous kidneys might be increased while maintaining an acceptable level of accuracy; this could have been used to avoid biopsies for six patients included in our study, without failing to perform any relevant biopsy.

We acknowledge the following limitations of our study. The classification of patients into "severe histopathologic changes" or "normal or mild histopathologic changes" according to their histopathologic findings is not an established dichotomy. Such classification was chosen because histopathologic findings of kidney transplants might include a broad range of pathologies; thus, it would be necessary to examine an inaccessible number of patients to detect DW-MRI patterns associated with specific pathologies. For this reason, in our study we focused on assessing for which patients the histopathologic findings had no implication on clinical management.

The employed reference standard might be biased since the biopsied specimens might not reflect the pathology of the entire kidney. There was a maximal time interval between MRI and biopsy of 10 days. It is not known how potential changes in therapy might have altered imaging or biopsy results; however, the median time interval between MRI and biopsy was 1 day. A further limitation is the retrospective design of the study and image analysis. The image analysis was performed in consensus and we did not assess the inter-reader variability when assigning kidneys to the homogeneous or heterogeneous group. In addition, we did not analyse the parameters' histogram over the whole kidney, which would have allowed an objective and purely quantitative classification of the patients; nevertheless, in clinical routine qualitative image analysis can be performed more quickly and easily. The precise indication criteria for each biopsy could not always be determined and not all kidney transplant biopsy reports were redacted according to the Banff classification, partly because of changes to the classification system during the study period [19]. The current Banff classification system, which is being used more and more by pathologists, also includes a quantitative assessment of histology findings. In the future, this may allow a direct correlation between DW-MRI and the severity of histopathologic findings.

Furthermore, diffusion in the kidney is highly anisotropic due to the presence of spatially oriented ducts with tubular flow. Therefore, it might be difficult to disentangle diffusion phenomena from vascular flow and consequently to obtain accurate estimates of diffusion and perfusion-related parameters [20]. Other MR techniques such as diffusion tensor imaging, arterial spin labelling, and blood oxygenation level-

Table 4 Results for 100% sensitivity

DW-MRI criteria for biopsy	Sensitivity (%)	Specificity (%)	Accuracy (%)	Corresponding cutoff value
Heterogeneous or homogeneous with a cutoff value for				
ADCmean [10^{-5} mm^2/s]	100.0 (25/25)	13.0 (2/15)	68.0 (27/40)	210
*D**mean [10^{-4} mm^2/s]	100.0 (25/25)	0.0 (0/15)	63.0 (25/40)	449
*D*mean [10^{-6} mm^2/s]	100.0 (25/25)	7.0 (1/15)	65.0 (26/40)	2070
*f*mean [%]	100.0 (25/25)	40.0 (6/15)	77.5 (31/40)	20
ADCmin [10^{-5} mm^2/s]	100.0 (25/25)	13.0 (2/15)	67.5 (27/40)	203
*D**min [10^{-4} mm^2/s]	100.0 (25/25)	0.0 (0/15)	62.5 (25/40)	326
*D*min [10^{-6} mm^2/s]	100.0 (25/25)	0.0 (0/15)	62.5 (25/40)	1950
*f*min [%]	100.0 (25/25)	40.0 (6/15)	77.5 (31/40)	16

dependent imaging might be promising functional imaging techniques for assessing structural histological changes in kidneys [21–27] and might allow assessing which kidney transplant patients need biopsy or even to forgo biopsies altogether.

In conclusion, the present study demonstrates that a combination of *qualitative* and *quantitative* DW-MRI allows predicting severe histopathologic alterations in a substantial number of kidney transplant patients without contrast medium administration. Q*ualitative* DW-MRI alone indicates the necessity of renal biopsy in heterogeneous renal allografts due to its high specificity. The low sensitivity based on qualitative image analysis can be further improved by adding quantitative analysis of homogeneous appearing kidney transplants.

Further prospective studies in a clinical setting are required to evaluate whether combined *qualitative* and *quantitative* DW-MRI can be used to predict the severity of histopathologic changes and the need for biopsies; however, in our opinion, functional MRI has the potential to reduce the amount of biopsies and perhaps one day even to replace the biopsies that are required today to diagnose patients with kidney allograft impairment.

Acknowledgements We would like to thank Peter Vermathen; PhD for setting up the MRI protocols.

Compliance with ethical standards

Guarantor The scientific guarantor of this publication is Harriet C. Thoeny.

Conflict of interest The authors of this manuscript declare no relationships with commercial companies.

Funding The authors state that this work has not received any funding.

Statistics and biometry S. Barbieri and M. Ith kindly provided statistical advice for this manuscript.

Both of the authors do have statistical expertise.

No complex statistical methods were necessary for this paper.

Ethical approval Institutional Review Board approval was obtained.

Informed consent Written informed consent was waived by the Institutional Review Board.

Methodology
- retrospective
- diagnostic or prognostic study/observational/experimental
- performed at one institution

References

1. Wilczek HE (1990) Percutaneous needle biopsy of the renal allograft. A clinical safety evaluation of 1129 biopsies. Transplantation 50:790–797
2. Furness PN, Philpott CM, Chorbadjian MT et al (2003) Protocol biopsy of the stable renal transplant: a multicenter study of methods and complication rates. Transplantation 76:969–973
3. Schwarz A, Gwinner W, Hiss M et al (2005) Safety and adequacy of renal transplant protocol biopsies. Am J Transplant 5:1992–1996
4. Group KDIGOKTW (2009) KDIGO clinical practice guideline for the care of kidney transplant recipients. Am J Transplant 9:S1–S155
5. Racusen LC, Solez K, Colvin RB et al (1999) The Banff 97 working classification of renal allograft pathology. Kidney Int 55:713–723
6. Nadalo LA, Dickerman RM, Slonim SM Imaging in kidney transplantation complications. http://emedicine.medscape.com/article/378801-overview. Accessed 16 Jul 2014
7. Thoeny HC, Zumstein D, Simon-Zoula S et al (2006) Functional evaluation of transplanted kidneys with diffusion-weighted and BOLD MR imaging : initial experience. Radiology 241:812–821
8. Eisenberger U, Thoeny HC, Binser T et al (2010) Evaluation of renal allograft function early after transplantation with diffusion-weighted MR imaging. Eur Radiol 20:1374–1383
9. Eisenberger U, Binser T, Thoeny HC et al (2014) Living renal allograft transplantation: diffusion-weighted MR imaging in longitudinal follow-up of the donated and the remaining kidney. Radiology 270:800–808
10. Abou-El-Ghar ME, El-Diasty TA, El-Assmy AM et al (2012) Role of diffusion-weighted MRI in diagnosis of acute renal allograft dysfunction: a prospective preliminary study. Br J Radiol 1–6
11. Marckmann P, Skov L, Rossen K et al (2006) Nephrogenic systemic fibrosis: suspected causative role of gadodiamide used for contrast-enhanced magnetic resonance imaging. J Am Soc Nephrol 17:2359–2362
12. Le Bihan D, Breton E, Lallemand D et al (1986) MR imaging of intravoxel incoherent motions: application to diffusion and perfusion in neurologic disorders. Radiology 161:401–407
13. Schindelin J, Rueden CT, Hiner MC et al (2015) The ImageJ ecosystem: an open platform for biomedical image analysis. Mol Reprod Dev 82:518–529
14. Baxter GM (2001) Ultrasound of renal transplantation. Clin Radiol 56:802–818
15. Charles JH (2007) Heptinstall's pathology of the kidney, 6th edn. Lippincott Wiliams & Wilkins, Philadelphia
16. Wu HH, Jia HR, Zhang Y, Liu L et al (2015) Monitoring the progression of renal fibrosis by T2-weighted signal intensity and diffusion weighted magnetic resonance imaging in cisplatin induced rat models. Chin Med J 128:626–631
17. Siegel M, Glanz S, Gordon DH, Butt KM (1981) Focal angiographic findings in renal transplant rejection. Urol Radiol 3:97–100
18. Kurtkoti J, Sakhuja V, Sud K et al (2008) The utility of 1- and 3-month protocol biopsies on renal allograft function: a randomized controlled study. Am J Transplant 8:317–323
19. Haas M, Sis B, Racusen LC et al (2014) Banff 2013 meeting report: inclusion of c4d-negative antibody-mediated rejection and antibody-associated arterial lesions. Am J Transplant 14:272–283
20. Lima M, Le Bihand D (2016) Clinical intravoxel incoherent motion and diffusion MR imaging: past, present, and future. Radiology 278:13–32
21. Blondin D, Lanzman RS, Mathys C et al (2009) Functional MRI of transplanted kidneys using diffusion-weighted imaging. Röfo 181:1162–1167
22. Lanzman RS, Wittsack H-J, Martirosian P et al (2010) Quantification of renal allograft perfusion using arterial spin labeling MRI: initial results. Eur Radiol 20:1485–1491
23. Lanzman RS, Ljimani A, Pentang G et al (2013) Kidney transplant: functional assessment with diffusion-tensor MR imaging at 3T. Radiology 266:218–225

24. Fan WJ, Ren T, Li Q et al (2016) Assessment of renal allograft function early after transplantation with isotropic resolution diffusion tensor imaging. Eur Radiol 26:567–575
25. Liu Z, Xu Y, Zhang J et al (2015) Chronic kidney disease: pathological and functional assessment with diffusion tensor imaging at 3T MR. Eur Radiol 25:625–660
26. Seif M, Eisenberger U, Binser T et al (2016) Renal blood oxygenation level-dependent imaging in longitudinal follow-up of donated and remaining kidneys. Radiology 279:795–804
27. Hueper K, Khalifa AA, Brasen JH et al (2016) Diffusion-Weighted imaging and diffusion tensor imaging detect delayed graft function and correlate with allograft fibrosis in patients early after kidney transplantation. J Magn Reson Imaging 44:112–121