

Received:
17 February 2017

Revised:
7 August 2017

Accepted:
30 August 2017

<https://doi.org/10.1259/bjr.20170125>

Cite this article as:

Hollis E, Shehata M, Abou El-Ghar M, Ghazal M, El-Diasty T, Merchant M, et al. Statistical analysis of ADCs and clinical biomarkers in detecting acute renal transplant rejection. *Br J Radiol* 2017; **90**: 20170125.

FULL PAPER

Statistical analysis of ADCs and clinical biomarkers in detecting acute renal transplant rejection

^{1,2}ELIZABETH HOLLIS, MSc, ¹MOHAMED SHEHATA, MSc, ³MOHAMED ABOU EL-GHAR, MD, ^{1,4}MOHAMMED GHAZAL, PhD, ³TAREK EL-DIASTY, MD, ²MICHAEL MERCHANT, PhD, ¹ANDREW E SWITALA and ¹AYMAN EL-BAZ, PhD

¹Department of Bioengineering, Bioluminescence Laboratory, University of Louisville, Louisville, KY, USA

²Department of Pharmacology and Toxicology, University of Louisville, Louisville, KY, USA

³Department of Radiology, Urology and Nephrology Center, University of Mansoura, Mansoura, Egypt

⁴Department of Electrical and Computer Engineering, Abu Dhabi University, Abu Dhabi, UAE

Address correspondence to: Dr Ayman El-Baz

E-mail: aselba01@louisville.edu

Elizabeth Hollis and Mohamed Shehata have contributed equally to this work.

Objective: The main goal of this study is to determine which parameters [e.g. clinical biomarkers, demographics and image-markers using 4D (3D + b -value) diffusion-weighted MRI (DW-MRI)] are more correlated with transplanted kidney status in patients who have undergone kidney transplantation, and can be used for early assessment of acute renal rejection.

Methods: The study included 16 patients with stable graft function and 37 patients with acute rejection (AR), determined by renal biopsy post-transplantation. 3D DW-MRI of each allograft had been acquired using a series of b -values 50 and 100–1000 in steps of 100 s mm⁻². The kidney was automatically segmented and co-aligned across series for motion correction using geometric deformable models. Volume-averaged apparent diffusion coefficients (ADCs) at each b -value were calculated. All possible subsets of ADC were used, along with patient age, sex, serum plasma creatinine (SPCr) and creatinine clearance (CrCl), as

predictors in 2¹¹ logistic regression models where AR was the outcome variable. Predictive value of ADC at each b -value was assessed using its Akaike weight.

Results: ANOVA of the saturated model found that odds of AR depended significantly on SPCr, CrCl and ADC at $b = 500, 600, 700$ and 900 s mm⁻². The model incorporating ADC at $b = 100$ and 700 s mm⁻² had the lowest value of the Akaike information criterion; the same two b -values also had the greatest Akaike weights. For comparison, the top 10 submodels and the full model were reported.

Conclusion: Preliminary findings suggest that ADC provides improved detection of AR than lab values alone. At least two non-zero gradient strengths should be used for optimal results.

Advances in knowledge: This paper investigated possible correlations between image-based and clinical biomarkers, and the fusion of both with respect to biopsy diagnosis of AR.

INTRODUCTION

The kidney is a significant organ playing vital roles in the maintenance of electrolytes, acid-base balance and blood pressure. It is the primary filtration organ in the human body, keeping in the nutrients that the body needs and expelling out the unwanted toxic wastes. A reduction of kidney function can occur naturally as a result of age or as a result of acute kidney injury or chronic disease (CKD). The primary causes of CKD can be attributed to diabetes and hypertension. However, other genetic (polycystic kidney disease or APOL1 nephropathy) or autoimmune (lupus nephritis or primary glomerular nephritis) diseases are additional causes for CKD.¹ While the primary insult may influence the rate of

CKD progression, the net effect of the unremitting accelerated loss of the kidney function is the build-up of uremic toxins in the body. The complete or near-complete loss of the kidney's ability to ultrafilter the blood and product urine is referred to as end-stage kidney disease (ESKD). Treatment of ESKD to address the deadly build-up of toxins is achieved through renal replacement therapy—dialysis or transplantation. Both of these approaches have inherent life altering complications. Hence, preserving the health of this organ is of the utmost importance.

The prevalence in the USA (2010–2014) of CKD was about 14.8% of the adult population and 680,000 cases of ESKD.

Owing to a limitation in available transplantable kidneys only approximately 17,000 yearly transplants were performed during this time.^{2,3} Although this has immensely improved the outcome of patients diagnosed with Stage 5 CKD (ESKD), there are still complications that can occur. One of the main concerns is graft dysfunction. Routine post-transplantation clinical evaluation of kidney function is of great importance to prevent the graft loss. The diagnostic technique presently recommended by the National Kidney Foundation to measure overall kidney function is glomerular filtration rate (GFR), which is based on measuring the serum creatinine level. However, this test has low sensitivity and is a late biomarker for renal dysfunction (a significant change in serum creatinine level is detectable only after the loss of 60% of renal function), and it does not assess the function of individual kidneys.⁴ The current gold standard for acute rejection (AR) diagnosis is needle biopsy. However, it is invasive, difficult to perform, costly and a time-consuming procedure. It can also result in over- or underestimates of AR by only sampling small kidney areas (*i.e.* the biopsy can underestimate AR if the sample size is small—not enough tissue pieces or not enough cortex/arteries in sample). Thus, the need for new non-invasive techniques to assess renal transplant status with the capability to provide accurate and early diagnosis of AR is of great clinical importance. A promising methodology that has been an area of increased research to diagnose graft dysfunction, without complications, is to utilize imaging techniques such as MRI. There are various types of MRI scans that are used for renal transplant assessment. While some of them provide only anatomical information, other MRI modalities, such as dynamic contrast-enhanced (DCE) and diffusion-weighted (DW) MRI, provide both anatomical and functional kidney information. An example for 2D DCE-MRI, Khalifa *et al*⁵ evaluated acute renal rejection post-transplantation. After segmenting transplanted kidneys, they measured two rates, namely, wash-in and wash-out rates that describe the perfusion of the contrast agent in the renal cortex in transient and plateau phases, and then, they interrelated those rates with the status of the transplanted kidney.

Clinically, nephrologists recommend the use of DCE-MRI if the GFR is greater than 30 ml min⁻¹ only, because the contrast agent may be nephrotoxic for those patients with a GFR less than 30 ml min⁻¹. To avoid the risk of nephrogenic systemic fibrosis and nephrotoxicity and to account for all GFR variability, DW-MRI is recommended. Therefore, we focused our study on the use of DW-MRIs in assessing renal transplants functionality. In fact, DW-MRI has become a subject of extensive research as an emerging imaging modality for renal function assessment thanks to DW-MRI's ability to provide both anatomical and functional information. For DW-MRI, its functional quantitative parameter, called apparent diffusion coefficient (ADC), is estimated from different gradient field strengths and durations (*b*-values) to describe the unique tissue characteristics of inner spatial water behaviour.⁶ Therefore, multiple studies^{7–15} have utilized DW-MRI to assess renal functionality by measuring the ADC values, but the results have been mixed and there is no agreement about specific ADC values at certain *b*-values that can differentiate AR from non-rejection (NR) renal transplants.⁷ In addition, most of these studies did not correlate estimated ADC values at

all *b*-values with both the clinical biomarkers and demographics of the patients. More details about these studies^{7–15} will be discussed later in the Discussion and Conclusions section.

Therefore, we were motivated to investigate which measurements are correlated with transplanted kidney status in patients who have undergone kidney transplantation, and can be used for early assessment of acute renal rejection by investigating the different renal allograft biomarkers; including: clinical biomarkers [*i.e.* serum plasma creatinine (SPCr) and creatinine clearance (CrCl)] alone; image-markers [*i.e.* the ADCs at the available 11 different individual *b*-values (*b*₅₀, and *b*₁₀₀ to *b*₁₀₀₀ with step of 100 s mm⁻²)] and the fusion of both the clinical biomarkers and the image-markers at different *b*-values.

MATERIALS

A total of 53 patients undergoing kidney transplantation provided consent to participate in this study with all scans and biopsies performed from July 2014 to June 2015. All kidney transplants were performed at the University of Mansoura, Egypt and the donated kidneys were from live donors. Patient characteristics including sex (44 males and 9 females); mean age 26.26 ± 9.87 years (range, 12–54 years). Patients were divided into two groups (Group 1 and Group 2). Group 1 (19 patients) included patients with healthy graft function. Most of Group 1 patients only underwent DW-MRI scans. However, five patients of Group 1 have some other symptoms, *i.e.* a significant change in the serum creatinine values even if they were in the normal range. This normal range is subjective and each patient has his own basal level usually of ≤1.3 mg dl⁻¹. These five patients were high risk to have rejection [*i.e.* false positive (F_p), which indicates the proportion of all negatives (NR) that still yield positive (rejection) test outcomes]; thus, they must perform the biopsy, but the biopsy confirmed that they have no rejection [*i.e.* true negative (T_N), which indicates the proportion of healthy people who are correctly identified as not having the rejection]. In Group 1, three patients were excluded from the study owing to technical problems (*e.g.* noisy artifacts during MRI scanning procedure coming from patient's movement) yielding total of 16 NR renal transplants.

Group 2 (41 patients) included patients with acute renal rejection, based on renal biopsy histology. In Group 2, four patients were excluded from the study owing to technical problems, yielding a total of 37 patients with rejected allografts. All Group 2 patients underwent DW-MRI and renal biopsy, which were performed together, respectively. Both DW-MRI and biopsy were included in the final analysis and examined by a nephrologist and a radiologist. Table 1 shows DW-MRI demographics and statistics. More details are documented in Appendix A.

MRI protocol

The MRI study was performed using a 1.5 T scanner (SIGNA horizon, General Electric Medical systems, Milwaukee, WI). DW-MR images were obtained by using a body coil and a single-shot spin-echo echo-planar sequence (TR/TE, 8000/61.2; bandwidth, 142 kHz; matrix, 1.28 × 1.28 mm²; section thickness, 4 mm; intersection gap, 0 mm; FOV, 36 cm; signals acquired, 7;

Table 1. Demographics statistics of the DW-MRI data for a total of 53 renal transplants included in the study

Variable	Overall (N = 53)	Non-rejected (N = 16)	Rejected (N = 37)
Age in years mean ($\pm$ SD)	26.2641509 ($\pm$ 9.87)	25.125 ($\pm$ 10.07)	26.7568 ($\pm$ 9.89)
Sex	Sex N (%)		
Male	44 (83%)	12 (75%)	32 (86%)
Female	9 (17%)	4 (25%)	5 (14%)
Age category (years)	Age category N (%)		
10–19	15 (28%)	5 (31%)	10 (27%)
20–29	19 (36%)	7 (44%)	12 (32%)
30–39	14 (26%)	2 (13%)	12 (32%)
40–49	4 (7%)	1 (7%)	2 (5%)
50–59	1 (2%)	1 (7%)	1 (3%)
Creatinine clearance (ml min ⁻¹)	Creatinine clearance N (%)		
10–19	3 (6%)	1 (6%)	2 (5%)
20–29	4 (8%)	0 (0%)	2 (5%)
30–39	4 (8%)	1 (6%)	3 (8%)
40–49	3 (6%)	1 (6%)	2 (5%)
50–59	6 (11%)	0 (0%)	6 (16%)
60–69	9 (17%)	0 (0%)	9 (24%)
70–79	13 (25%)	7 (44%)	6 (16%)
80–89	6 (11%)	2 (4%)	4 (11%)
90–99	2 (4%)	1 (6%)	1 (3%)
100–109	2 (4%)	1 (6%)	1 (3%)
110–119	2 (4%)	2 (4%)	0 (0%)
Serum plasma creatinine (mg dl ⁻¹)	Serum plasma creatinine N (%)		
0.0–0.9	4 (8%)	3 (19%)	1 (3%)
1.0–1.9	37 (70%)	10 (63%)	27 (73%)
2.0–2.9	8 (15%)	2 (13%)	6 (16%)
3.0–3.9	1 (2%)	0 (0%)	1 (3%)
4.0–4.9	2 (4%)	0 (0%)	2 (5%)
5.0–5.9	1 (2%)	1 (6%)	0 (0%)

Demographics include; sex, age, serum creatinine clearance and plasma creatinine. Standard deviation indicated by ($\pm$ SD), number of patients in category indicated by *N*, percentage of patients in the diagnosis (*i.e.* overall, non-rejected and rejected) indicated by (%).

water signals acquired at different *b*-values of (*b*₀, *b*₅₀, *b*₁₀₀, *b*₂₀₀, *b*₃₀₀, *b*₄₀₀, *b*₅₀₀, *b*₆₀₀, *b*₇₀₀, *b*₈₀₀, *b*₉₀₀ and *b*₁₀₀₀) s mm⁻². Approximately 50 sections have been obtained in 60–120 s to cover the whole kidney pelvis. Figure 1 shows a sample 4D (3D + *b*-value) DW-MRI scan for one subject.

METHODS AND MATERIALS

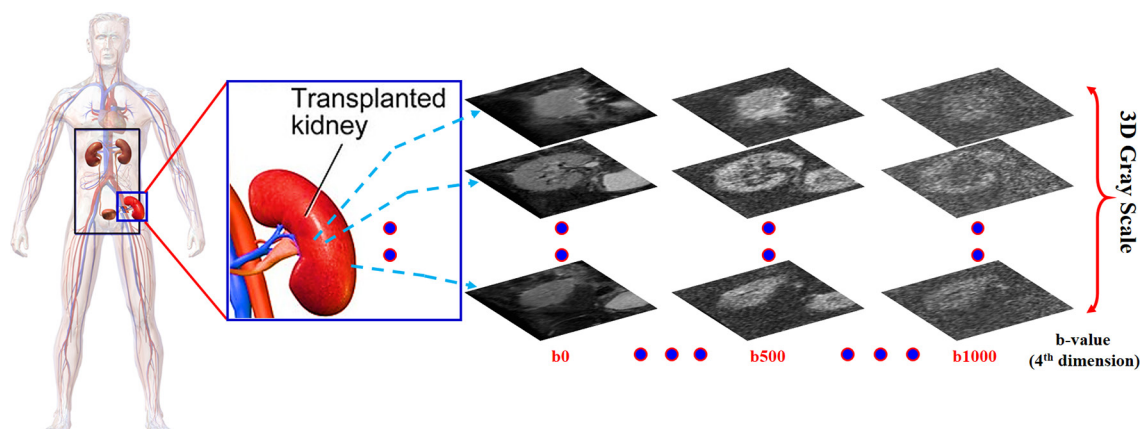
The primary goal of this retrospective pilot study was to determine which image-based measurements, derived from 4D (3D + *b*-value) DW-MRI, and/or clinical-based measurements are correlated with AR, which can be used for a non-invasive early assessment of AR post-transplantation. To this end, all patients' kidneys were evaluated using: (i) DW-MRI physiological biomarkers alone; namely, ADCs measured at different *b*-values;

(ii) clinical biomarkers [*e.g.* SPCr and CrCl] and patients' demographics (*i.e.* age and sex) and (iii) integration of both (i) and (ii). Then, statistical analysis was performed to investigate possible correlations between renal allograft biomarkers and the biopsy diagnosis of either AR or NR renal transplant. Details are outlined below.

DW-MR image analysis

An important and significant advantage of DW-MRI is the ability to interrelate local blood diffusion characteristics with the transplant status. This advantage is achieved through the DW-MRIs ability to measure unique tissue characteristics of inner spatial water behaviour called ADC,¹⁶ which can be used to assess the transplant status. In order to obtain accurate estimation of

Figure 1. Illustration for the 4D DW-MRI (3D + b value) scan for a transplanted kidney, where the 3D comes from acquiring several coronal cross sections to cover the whole kidney volume and the fourth dimension is coming from acquiring the same cross sections at different gradient field strength and duration known as b -values.



the DW-MRI-derived markers, namely ADCs, multiple image processing steps have to be performed using our previously developed approach.^{17,18} First, the noise effects and image inhomogeneity were reduced for a given DW-MRI data by applying an intensity histogram equalization and the non-parametric bias correction technique.¹⁹ This was followed by a 3D B-splines-based non-rigid registration to handle kidney motion to reduce the DW-MRI data variability across subjects.²⁰ Then, the kidney was segmented using a geometric deformable model guided by a joint Markov-Gibbs random field model accounting for both kidney shape prior information and kidney/background intensity appearance.^{21–23} After segmenting kidneys, the ADCs were estimated at different gradient field strengths and durations (b -values, in our case 11 b -values) using the equation defined by Le Bihan²⁴ as:

$$ADC_p = \frac{1}{b_0 - b} \ln \left(\frac{g_{b:p}}{g_{0:p}} \right)$$

such that $p = (x; y; z)$ denotes a voxel at a position with discrete Cartesian coordinates $(x; y; z)$, g_0 represents the segmented DW-MR image acquired at the b_0 and g_b represents the DW-MR image acquired at a given different b -value. Figure 2 illustrates voxel-wise ADC maps calculations for a segmented kidney subject.

Furthermore, colour maps were depicted to visually demonstrate the local voxel-wise diffusion of the segmented DW-MRI data. These regional display mappings are of great importance for the radiologists to help investigate which region of the kidney needs attention and follow-up with appropriate treatment. Figure 3 and 4 demonstrate the voxel-wise parametric maps for the diffusion of the transplanted kidney for a NR case and an AR case, respectively. The data in Figure 3 and 4 reveals the expected relation of the DW-MRI parameters for NR vs. AR status.

Statistical analysis

After the DW-MRI-derived markers were estimated (*i.e.* ADCs), statistical analysis was performed to investigate possible correlations between renal allograft biomarkers (clinical-based and image-based) and the biopsy diagnosis. The statistical

analysis examined four categories of parameters: (i) clinical biomarkers [*i.e.* SPCr and CrCl] alone, (ii) the mean ADC at 11 different individual b -values [b_{50} , and b_{100} to b_{1000} with step of 100 s mm^{-2}] (iii) all possible combinations of the mean ADCs of individual b -values (*i.e.* 2^{11} submodels) and (iv) the fusion of the clinical biomarkers with the mean ADC of fused b -values [the full model].

Statistical calculations were performed using R v. 3.1.1 (R Foundation, Vienna, Austria). The relationship of graft tolerance to ADC was tested using logistic regression of biopsy result (AR or NR) against demographic, laboratory and imaging parameters. The full model is the model that included age and sex, SPCr and CrCl and mean ADC at each b -value. Statistical significance of each parameter in the full model was assessed using likelihood ratio (χ^2) tests. Reduced logistic regression models were fit to the data, where each reduced model included demographics and laboratory variables along with a subset of mean ADC values. Considering all possible subsets results in 2^{11} models being fit, including the full model and the “null model” with demographics and lab results only. Identification of “most informative” b -values was performed using Akaike information criterion (AIC) using the following equation as:

$$AIC = -2l + 2m$$

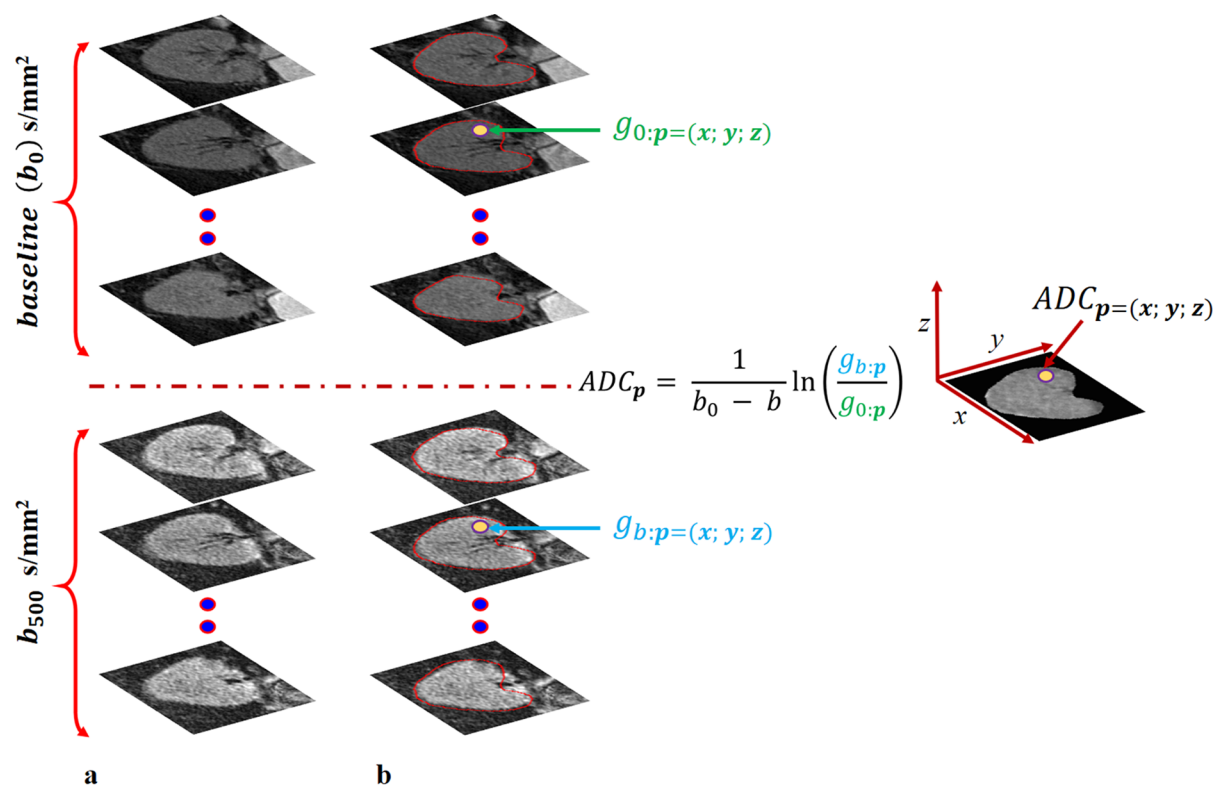
Where m is the number of free parameters in the model, and l is the log-likelihood of the fitted model when all free parameters have been adjusted to maximize the likelihood.

Information conferred by each mean ADC measure was quantified by the cumulative Akaike weight of models where it was included. Let $\Delta_i = AIC_i - \min_j AIC_j$ for $i, j = 1, \dots, 2048$, and $w_i = e^{-\Delta_i/2} / \sum_{j=1}^{2048} e^{-\Delta_j/2}$. Then, the relative importance of variable ADC b is

$$I_b = \sum_{i \in M} w_i,$$

where M is the set of indices of models including ADC b .

Figure 2. Illustration sample for a segmented kidney subject, where (a) the raw data at b_0 and b_{500} s mm^{-2} , (b) the segmented kidney object at b_0 and b_{500} s mm^{-2} and (c) the voxel-wise ADC maps calculation across the subject.



However, some questions must be answered regarding the statistical significant difference between the NR and the AR renal transplants using: (i) the clinical biomarkers (*i.e.* SPCR

and CrCl) alone, (ii) the mean ADCs at different individual 11 b -values from (b_{50} to b_{1000}). In addition, two more questions need to be answered regarding how informative the built

Figure 3. Illustration sample for a non-rejection subject, where (a) the raw data at the b -value of 0 s mm^{-2} , (b) the segmented kidney object at the b_0 s mm^{-2} and (c) the average voxel-wise diffusion parametric maps across the subject.

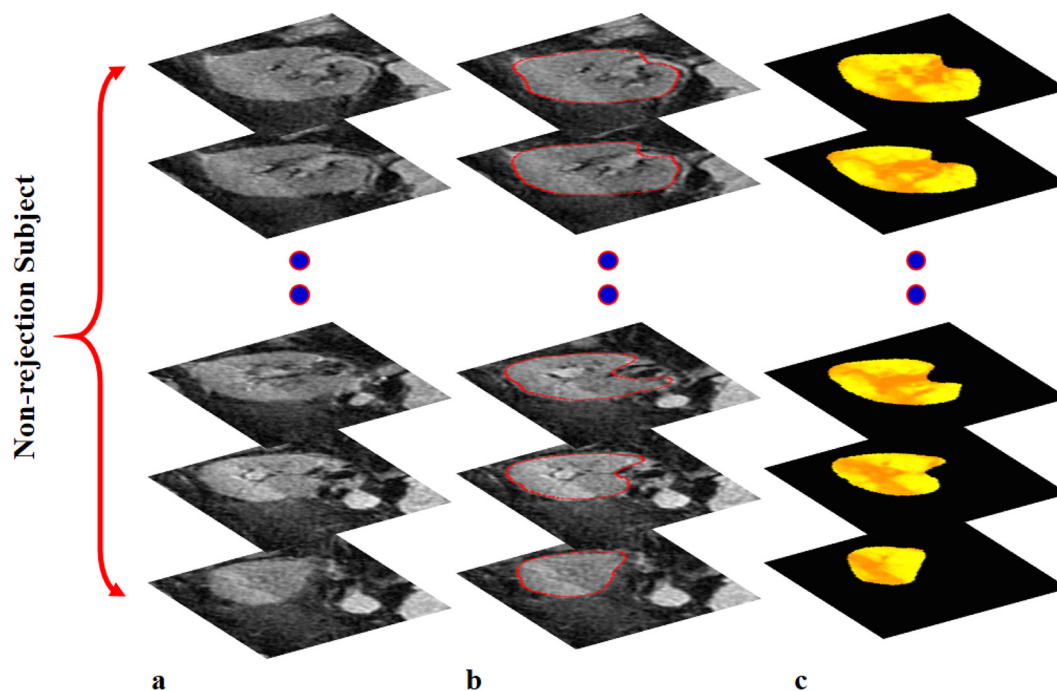
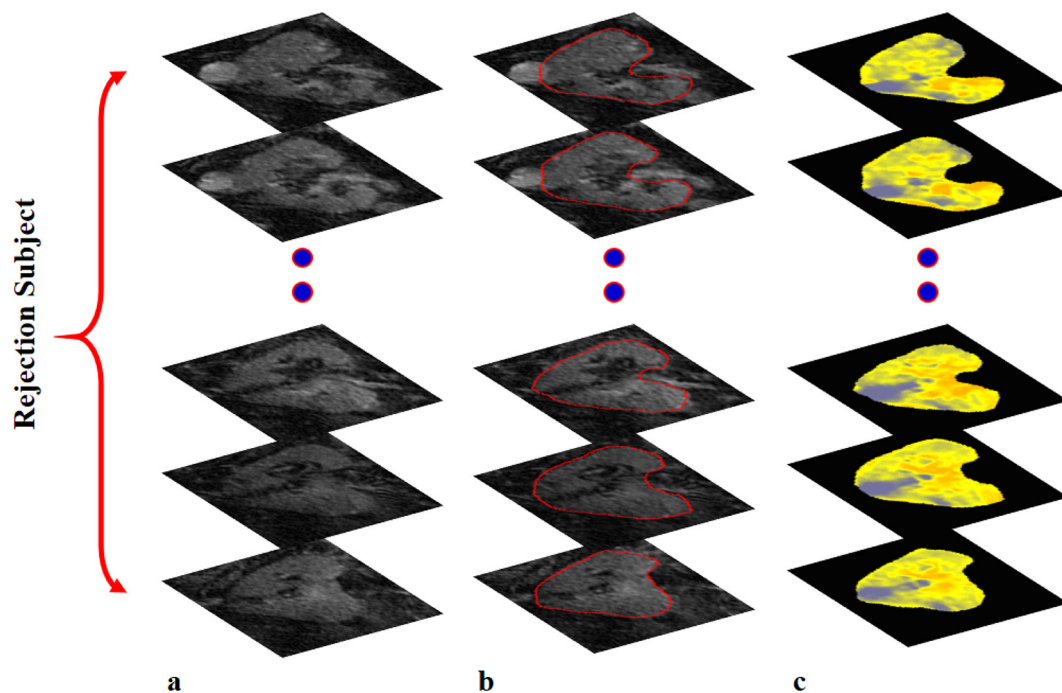


Figure 4. Illustration sample for a rejection subject, where (a) the raw data at the b -value of 0 s mm^{-2} , (b) the segmented kidney object at the $b_0 \text{ s mm}^{-2}$ and (c) the average voxel-wise diffusion parametric maps across the subject.



model is using: (i) the mean ADCs of a certain group of individual b -values (submodel) and (ii) the fusion of the clinical biomarkers with the mean ADCs of fused b -values (the full model).

RESULTS

ANOVA of the full model found that SPCr ($\chi^2 = 10.1$, $p = 0.002$) and CrCl ($\chi^2 = 14.1$, $p = 0.0002$) had a significant effect on the likelihood of AR, as did the mean ADC for $b = 500 \text{ s mm}^{-2}$ ($\chi^2 = 3.98$, $p = 0.0461$), $b = 600 \text{ s mm}^{-2}$ ($\chi^2 = 5.81$, $p = 0.0159$), $b = 700 \text{ s mm}^{-2}$ ($\chi^2 = 5.65$, $p = 0.0174$) and $b = 900 \text{ s mm}^{-2}$ ($\chi^2 = 4.94$, $p = 0.0262$). Patient age, sex and mean ADC at other b -values were not statistically significant, as shown in Table 2.

In addition, the mean ADC values and their associated SD at each b -value for both groups; Group 1 (NR) and Group 2 (AR) are documented in Table 3. Furthermore, to investigate which threshold ADC value indicates AR, a linear discriminant analysis was made at the four-different b -values, namely b_{500} , b_{600} , b_{700} and $b_{900} \text{ s mm}^{-2}$, of which the mean ADCs have showed statistically significant effect on the likelihood of AR. These values are also reported in Table 3. Furthermore, the same linear discriminant analysis was made for the laboratory tests, namely SPCr and CrCl, which have been found to be statistically significant as well, to find the threshold values of both SPCr and CrCl, which indicate AR as well. These threshold values of SPCr and CrCl were found to be 66.6 mg dl^{-1} and 1.66 ml min^{-1} , respectively, as documented in Table 3.

In addition, the identification of “most informative” b -values was performed using AIC as discussed in the Method Section. The reduced model (RM₁) incorporating only ADC at $b = 100 \text{ s mm}^{-2}$

and at $b = 700 \text{ s mm}^{-2}$ had the lowest AIC = 58.6, as shown in Table 4, and mean ADC at these same two b -values were the most informative predictors of AR according to their summed Akaike

Table 2. Results of analysis of covariance on the logistic regression model for acute rejection

Variable	c	χ^2	p -value
ADC ₅₀	-4.5 s mm^{-2}	2.618	0.106
ADC ₁₀₀	15.0 s mm^{-2}	3.791	0.0515
ADC ₂₀₀	-11.3 s mm^{-2}	0.610	0.435
ADC ₃₀₀	43.8 s mm^{-2}	1.242	0.265
ADC ₄₀₀	-8.8 s mm^{-2}	0.016	0.899
ADC ₅₀₀	-158.1 s mm^{-2}	3.978	0.0461
ADC ₆₀₀	266.1 s mm^{-2}	5.813	0.0159
ADC ₇₀₀	-201.8 s mm^{-2}	5.654	0.0174
ADC ₈₀₀	181.3 s mm^{-2}	1.446	0.229
ADC ₉₀₀	-122.0 s mm^{-2}	4.944	0.0262
ADC ₁₀₀₀	-52.7 s mm^{-2}	0.087	0.769
Age	0.065 y^{-1}	0.846	0.358
Sex = M	1.2	2.291	0.130
CrCl	-0.2 min ml^{-1}	14.068	0.000176
SPCr	-4.3 dl mg^{-1}	10.070	0.00151

ADC, apparent diffusion coefficient; ADC_{*n*}, mean ADC at $b = n$; c , fitted coefficient (log-odds ratio); CrCl, creatinine clearance; SPCr, serum plasma creatinine.

The highlighted bold type represents statistically significant b -values and CrCl and SPCr and that they have a significance p -value < 0.05 .

Table 3. Mean CrCl value, mean SPCr value, their associated SDs, their threshold values that indicate AR renal transplants, mean ADC values, their associated SDs at each b -value from b_{50} to b_{1000} s mm^{-2} for both AR and NR renal transplants and the threshold ADC values that indicate AR at the statistical significant b -values only; namely, b_{500} , b_{600} , b_{700} and b_{900} s mm^{-2}

Lab tests	AR mean	AR (SD)	NR mean	NR (SD)	Threshold values
CrCl ml min^{-1}	59.4	(19.8)	73.8	(26.5)	66.6
SPCr mg dl^{-1}	1.78	(0.85)	1.54	(1.14)	1.66
b -value (s mm^{-2})	ADC ($\text{mm}^2 \text{s}^{-1}$)				
	AR mean	AR (SD)	NR mean	NR (SD)	Threshold ADC values
b_{50}	4.66	(0.75)	4.53	(0.53)	–
b_{100}	2.33	(0.37)	2.27	(0.21)	–
b_{200}	1.16	(0.14)	1.16	(0.11)	–
b_{300}	0.805	(0.085)	0.816	(0.069)	–
b_{400}	0.603	(0.053)	0.616	(0.049)	–
b_{500}	0.526	(0.044)	0.542	(0.042)	0.534
b_{600}	0.417	(0.034)	0.427	(0.028)	0.422
b_{700}	0.374	(0.029)	0.394	(0.031)	0.384
b_{800}	0.286	(0.023)	0.295	(0.026)	–
b_{900}	0.241	(0.021)	0.250	(0.024)	0.245
b_{1000}	0.221	(0.019)	0.230	(0.023)	–

weighting (Figure 5). For comparison, the full model, *i.e.* the model that includes all the mean ADCs from all b -values with demographics (age and sex) and clinical biomarkers (CrCl and SPCr), has been included with AIC = 65.0.

DISCUSSION AND CONCLUSIONS

Using 4D (3D + b -value) DW-MRI, we determined the statistical significance of measured parameters [clinical biomarkers,

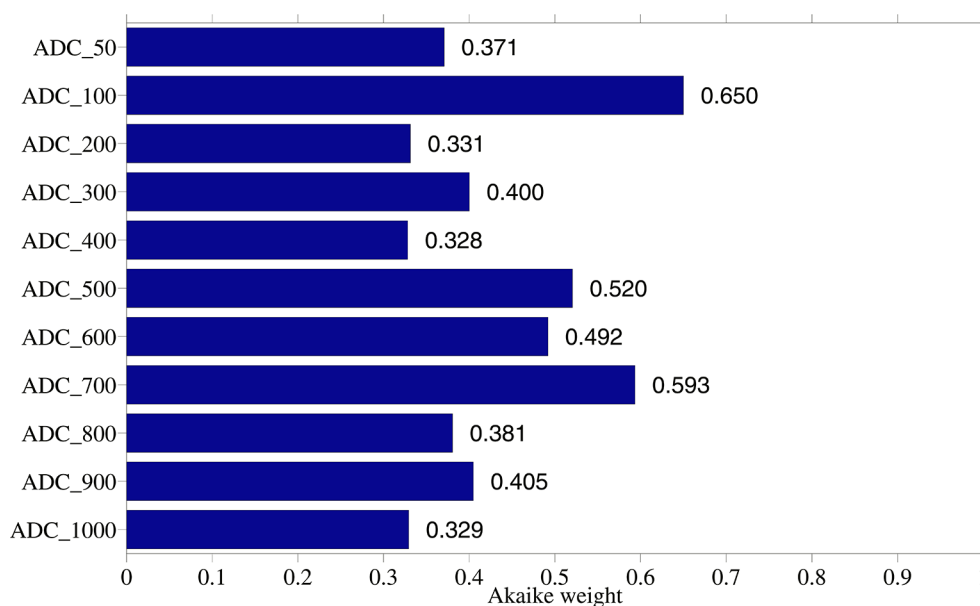
mean ADC at different individual b -values (b_{50} to b_{1000}), submodels and full model] that might be helpful in detecting acute renal rejection in patients who had undergone kidney transplantation. We hypothesized the most accurate parameters for diagnosing would be the full model, owing to the fact that all the mean ADC of each b -value would have been fused giving all the information of the scan and, therefore, giving the whole picture. To the best of our knowledge, this is the first

Table 4. Mean ADC included in the models with the 10 lowest AIC values, these are indicated by reduced model (RM_i), $i = 1, \dots, 10$

Model	AIC	Δ	Predictor variables (b -value)										
			50	100	200	300	400	500	600	700	800	900	1000
RM_1	58.6 (min)	0											
RM_2	58.7	0.04											
RM_3	59.2	0.57											
RM_4	59.6	0.96											
RM_5	59.7	1.04											
RM_6	59.8	1.14											
RM_7	59.8	1.16											
RM_8	59.9	1.20											
RM_9	59.9	1.28											
RM_{10}	60.0	1.39											
DG	61.9	3.29											
Full	65.0	6.40											

The full model (Full), *i.e.* the model that incorporates all the mean ADCs from all b -values with demographics (age and sex) and clinical biomarkers (CrCl and SPCr), and the DG model, with only demographics (age and sex) and clinical biomarkers (CrCl and SPCr), are shown for comparison. $\Delta = \text{AIC} - \text{min AIC}$, where min AIC is RM_1 with the lowest AIC value as indicated in the table.

Figure 5. The Akaike weighting criterion vs. the submodels at different b -values, reflects how much informative each submodel is.



systematic study of the best b -values for kidney DW-MRI in the range 50–1000 s mm⁻². ANOVA of the full model found that the clinical biomarkers (SPCr and CrCl) had a significant effect on the likelihood of AR, which conform with the results of the studies made in Eisenberger et al¹¹ and Wypych-Klunder et al.¹⁴ Additionally, image markers (the mean ADCs for $b = 500$, $b = 600$, $b = 700$ and $b = 900$ s mm⁻²) were found to have a statistical significant effect on the likelihood of AR as well, which is consistent with the results obtained in Hueper et al,¹⁰ Eisenberger et al,¹¹ Abou-El-Ghar et al¹³ and Wypych-Klunder et al.¹⁴ Based on the AIC, the most informative model with the most accurate parameters would be the submodel that fuses the image markers (mean ADC at $b = 100$ s mm⁻² and $b = 700$ s mm⁻²) with the clinical biomarkers (CrCl and SPCr) and demographics (age and sex) and can be used in assessing renal transplant status. All other parameter's accuracy lies in between this submodel and the full model. On the other hand, patient age, sex and mean ADC at other b -values were not statistically significant. Also, it was clear that the NR group has higher mean ADC values than AR group, at least for $b \geq 300$ s mm⁻² (Table 4), which conform with the results obtained from.^{7–15} However, each b -value has its own threshold ADC value that might be useful in deafferenting AR from NR kidney transplants. This can be explained in part by that one of the most significant advantage of DW-MRI is its ability to interrelate local blood diffusion characteristics with the transplant status. This advantage is achieved through the DW-MRIs ability to measure unique tissue characteristics of inner spatial water behaviour (*i.e.* how freely water can move within a voxel) known as ADC, which were investigated to be higher in the NR group than the AR group. That means AR renal transplants explain lower diffusivity than NR renal transplants because of the existence of some local tissue regions that have bad blood diffusion. Compared with our findings, some of the existing studies that used DW-MRI in diagnosing AR renal transplants are briefly discussed.

A study by Liu et al⁷ explored the detection of early renal allograft dysfunction caused by AR using DW-MRI. With manually selected medullary and cortical regions of interests (ROIs), lower ADC values of the AR group than those of the control group were revealed. A similar earlier study was conducted by Thoeny et al⁸ and Xu et al⁹ investigated the potential power of DW-MRI to diagnose AR renal allografts on 26 biopsy-proven rejection and 43 NR patients. They found that higher ADC values were obtained from the normal allografts compared with those from AR allografts. The receiver operating characteristic curve was constructed and demonstrated the best sensitivity and specificity at the b -value of 800 s mm⁻². A study by Hueper et al¹⁰ was conducted to assess renal allografts functionality included 64 patients with renal allografts, of which 33 were patients with initial graft function and 31 were patients with delayed graft function. These patients underwent DW-MRI scans at two b -values (0 and 600 s mm⁻²). After placement of manual ROIs in the lower, middle and upper poles of the medulla and cortex on several portions to cover large regions of the allograft and estimation of renal diffusion parameters, including ADC, they concluded that renal diffusion parameters were significantly reduced in patients with delayed graft function and their values well correlated with renal function in biopsy specimens. Likewise, a recent study by Eisenberger et al¹¹ was also conducted to evaluate renal allograft functionality. They began with a manual placement of ROIs. Then, they calculated the means and SDs of the ADCs from all b -values. A significant reduction in these parameters was observed in the cortex and the medulla for the AR patients, and the previously stated parameters were correlated with the CrCl values. Kaul et al¹² examined renal dysfunction assessment using cortical and medullary ADC maps. They found a significant decrease in ADC values of medullas compared with those cortices in normal donor kidneys and normal allografts. Both the cortex and medulla ADCs decreased or increased significantly for the

the rejection or the recovery from the rejection itself, respectively. Abou El-Ghar et al¹³ assessed renal functionality for 70 renal allograft patients. DW-MRI scans at two b -values of 0 and 800 s mm⁻² were performed for 49 patients with normal renal allografts (Group 1) and 21 patients with acute graft impairment (Group 2). In a single cross-section, a user-defined ROI was placed in the middle of the kidney and a pixel-wise ADC was calculated. Results show that the ADC values of Group 2 were significantly lower than those of Group 1. Possible relations between the selected laboratory results and diffusion parameters in the early period post kidney transplantation was explored by Katarzyna et al.¹⁴ These measurements were conducted in the kidney's cortex and medulla over multiple user-defined ROIs at b_{600} and b_{1000} s mm⁻². They obtained the best-quality ADC measurement in the renal cortex at the b -value of 1000 s mm⁻² because of the relative variability of results and signal-to-noise ratio. In addition, strong dependencies were observed between the ADC and exponential ADC, measured in the renal cortex at b_{1000} s mm⁻², and the estimated GFR. Vermathen et al¹⁵ inspected the determination of long-term (3 years) stability and potential changes for renal allograft recipients. Cortical and medullary ROIs were designated and the ADC were measured from all b -values. A significant correlation between different ADC components was demonstrated in the case of normal transplants.

Although our preliminary results of the proposed technique hold great promise in diagnosing renal graft dysfunction, there are several limitations with this study that we plan to overcome in a bigger study in the future. Long-term follow-up data were not available, so it is unknown whether any of the patients who appeared to tolerate the graft at the initial biopsy later developed AR syndrome. The greatest limitation of this pilot study was its small sample size; which if increased might affect the statistical significant b -values, especially those on the edge of significance (e.g. b_{100} s mm⁻² with p -value = 0.0515); thus, we plan to test the proposed technique on a larger and more diverse cohort of patients to confirm the accuracy and robustness of the proposed technique. This brings us to another limitation of this study; the diversity of the sample was limited. This study was performed in Egypt and, therefore, was limited to one race of the sample of patients that were in the study. Therefore, we are currently running another study at different US sites (e.g. the University of Louisville and the University of Michigan). It would be

appropriate to have a sample that includes different races and/or cultural background to reflect a wider population of kidney transplant patients. Also, as one can see from Table 1, there is an overwhelming number of males in this study. Although this study does demonstrate in Table 2 that there is no significant difference between males and females, the findings of the study may not accurately portray the ability of the proposed technique to diagnose renal graft dysfunction in females. It would be of the next studies best interest to have a sample size that is $\geq 50\%$. By correcting this, the female population with a kidney transplant can be better diagnosed in the future. One more limitation of this study is the age of the sample size, although it does cover a wide range of ages, it does exclude patients that are over 60 and under 10 years of age. In addition, new data set with lower b -values, which describes blood perfusion, will be included to investigate whether these lower b -values will be more helpful in finding more physiologically significant differences between NR and rejection transplants.

In conclusion, our study demonstrated that by fusing image markers derived from specific b -values, namely ADCs, with the clinical biomarkers and demographics, one can get a more accurate picture than if one were to use clinical biomarkers and demographics alone, ADC from a single b -value alone, or fused ADCs from all b -values together without clinical biomarkers or demographics. Our analysis has shown that the proposed technique utilizing DW-MR images (namely, ADC at b_{100} and b_{700} s mm⁻²) along with the clinical biomarkers (SPCr and CrCl) and demographics (age and sex), holds a promise as a reliable non-invasive diagnostic tool. In addition, our analysis demonstrated higher ADC values of NR than those of AR renal transplants. As mentioned above, we are currently collecting new data at the University of Louisville and the University of Michigan to conduct a bigger study with a larger number, more diverse and more gender balanced participating renal transplants. So far, this study only focuses on detecting the global transplant status as NR or AR. The ability of the developed technique needs to be examined in diagnosing different types of rejection (e.g. T-cell and antibody mediated rejection) for proper treatment administration.

FUNDING

This research is supported by the US-Egypt Science and Technology (S&T) Joint Fund #2000007145.

REFERENCES

1. Kukla A, Adulla M, Pascual J, Samaniego M, Nanovic L, Becker BN, et al. CKD stage-to-stage progression in native and transplant kidney disease. *Nephrol Dial Transplant* 2008; 23: 693–700. doi: <https://doi.org/10.1093/ndt/gfm590>
2. Kidney.org. About Chronic Kidney Disease. 2016. Available from: <https://www.kidney.org/kidneydisease/aboutckd> [cited 2 May 2016].
3. National Kidney Foundation. Organ Donation and Transplantation Statistics. 2016. Available from: <https://www.kidney.org/news/newsroom/factsheets/Organ-Donation-and-Transplantation-Stats> [cited 25 May 2016].
4. Myers GL, Miller WG, Coresh J, Fleming J, Greenberg N, Greene T, et al. Recommendations for improving serum creatinine measurement: a report from the laboratory working group of the national kidney disease education program. *Clin Chem* 2006; 52: 5–18. doi: <https://doi.org/10.1373/clinchem.2005.0525144>
5. Khalifa F, Abou El-Ghar M, Abdollahi B, Frieboes HB, El-Diasty T, El-Baz A. A comprehensive non-invasive framework for automated evaluation of acute renal transplant rejection using DCE-MRI. *NMR*

- Biomed* 2013; **26**: 1460–70. doi: <https://doi.org/10.1002/nbm.2977>
6. Michaely HJ, Herrmann KA, Nael K, Oesingmann N, Reiser MF, Schoenberg SO. Functional renal imaging: nonvascular renal disease. *Abdom Imaging* 2007; **32**: 1–16. doi: <https://doi.org/10.1007/s00261-005-8004-0>
 7. Liu G, Han F, Xiao W, Wang Q, Xu Y, Chen J. Detection of renal allograft rejection using blood oxygen level-dependent and diffusion weighted magnetic resonance imaging: a retrospective study. *BMC Nephrol* 2014; **15**: 158. doi: <https://doi.org/10.1186/1471-2369-15-158>
 8. Thoeny HC, Zumstein D, Simon-Zoula S, Eisenberger U, De Keyser F, Hofmann L, et al. Functional evaluation of transplanted kidneys with diffusion-weighted and BOLD MR imaging: initial experience. *Radiology* 2006; **241**: 812–21. doi: <https://doi.org/10.1148/radiol.2413060103>
 9. Xu JJ, Xiao WB, Zhang L, Zhang MM. Value of diffusion-weighted MR imaging in diagnosis of acute rejection after renal transplantation. *Zhejiang Da Xue Xue Bao Yi Xue Ban* 2010; **39**: 163–7.
 10. Hueper K, Khalifa AA, Bräsen JH, Vo Chieu VD, Gutberlet M, Wintterle S, et al. 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* 2016; **44**: 112–21. doi: <https://doi.org/10.1002/jmri.25158>
 11. Eisenberger U, Thoeny HC, Binser T, Gugger M, Frey FJ, Boesch C, et al. Evaluation of renal allograft function early after transplantation with diffusion-weighted MR imaging. *Eur Radiol* 2010; **20**: 1374–83. doi: <https://doi.org/10.1007/s00330-009-1679-9>
 12. Kaul A, Sharma RK, Gupta RK, Lal H, Yadav A, Bhadhuria D, et al. Jaisuresh Assessment of allograft function using diffusion-weighted magnetic resonance imaging in kidney transplant patients. *Saudi J Kidney Dis Transpl* 2014; **25**: 1143. doi: <https://doi.org/10.4103/1319-2442.144245>
 13. Abou-El-Ghar ME, El-Diasty TA, El-Assmy AM, Refaie HF, Refaie AF, Ghoneim MA. Role of diffusion-weighted MRI in diagnosis of acute renal allograft dysfunction: a prospective preliminary study. *Br J Radiol* 2014;.
 14. Wypych-Klunder K, Adamowicz A, Lemanowicz A, Szczęśny W, Włodarczyk Z, Serafin Z. Diffusion-weighted MR imaging of transplanted kidneys: preliminary report. *Pol J Radiol* 2014; **79**: 94–8. doi: <https://doi.org/10.12659/PJR.890502>
 15. Vermathen P, Binser T, Boesch C, Eisenberger U, Thoeny HC. Three-year follow-up of human transplanted kidneys by diffusion-weighted MRI and blood oxygenation level-dependent imaging. *J Magn Reson Imaging* 2012; **35**: 1133–8. doi: <https://doi.org/10.1002/jmri.23537>
 16. Hollis E, Shehata M, Khalifa F, Abou El-Ghar M, El-Diasty T, El-Baz A. Towards non-invasive diagnostic techniques for early detection of acute renal transplant rejection: A review. *Egypt J Radiol Nucl Med* 2017; **48**: 257–69. doi: <https://doi.org/10.1016/j.ejrnrm.2016.11.005>
 17. Shehata M, Khalifa F, Soliman A, Takieldean A, El-Ghar MA, Shaffie A, et al. 3D diffusion MRI-based CAD system for early diagnosis of acute renal rejection. In: *Proceedings of IEEE 13th International Symposium on Biomedical Imaging*, 2016. New Jersey, USA: IEEE; 2016. pp. 1177–80.
 18. Shehata M, Khalifa F, Soliman A, El-Ghar MA, Dwyer A, Gimelfarb G, et al. A promising non-invasive CAD system for kidney function assessment. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention* 2016 Oct 17. Cham, Switzerland: Springer International Publishing; 2016. pp. 613–21.
 19. Tustison NJ, Avants BB, Cook PA, Zheng Y, Egan A, Yushkevich PA, et al. N4ITK: improved N3 bias correction. *IEEE Trans Med Imaging* 2010; **29**: 1310–20. doi: <https://doi.org/10.1109/TMI.2010.2046908>
 20. Glocker B, Komodakis N, Paragios N, Navab N. Non-rigid registration using discrete MRFs: application to thoracic CT images. In: *Workshop Evaluation of Methods for Pulmonary Image Registration in conjunction with Medical Image Computing and Computer-Assisted Intervention*. Springer International Publishing; 2010.
 21. Shehata M, Khalifa F, Soliman A, Alrefai R, El-Ghar MA, Dwyer AC, et al. A novel framework for automatic segmentation of kidney from DW-MRI. In: *Proceedings of IEEE 12th International Symposium on Biomedical Imaging*, 2015. New Jersey, USA: IEEE; 2015. pp. 951–4.
 22. Shehata M, Khalifa F, Soliman A, Alrefai R, El-Ghar MA, Dwyer AC, et al. A level set-based framework for 3D kidney segmentation from diffusion MR images. In: *Proceedings of IEEE 22nd International Conference on Image Processing*, 2015. New Jersey, USA: IEEE; 2015. pp. 4441–5.
 23. El-Baz A, Jiang X, Suri J. Advances & trends. In: *Biomedical image segmentation*. Boca Raton, Florida: CRC Press; 2016.
 24. Le Bihan D, Breton E. Imagerie de diffusion in-vivo par résonance magnétique nucléaire. *C R Acad Sci* 1985; **93**: 27–34.

APPENDIX A: DW-MRI DATA AND DIAGNOSIS

Patient number	Demographics		Laboratory tests		Biopsy findings	
	Gender	Age	Creatinine clearance (ml min ⁻¹)	Plasma creatinine (mg dl ⁻¹)	Biopsy diagnosis	Banff scores
Patient #001	F	15	70	1.1	NR	–
Patient #002	F	26	70	0.9	NR	–
Patient #003	M	14	80	1.1	NR	–
Patient #004	M	17	80	0.9	NR	–
Patient #005	M	44	96	1.1	NR	–
Patient #006	M	36	111	1.1	NR	–
Patient #007	M	18	112	1.1	NR	–
Patient #008	M	20	102	0.8	NR	–
Patient #009	M	22	75	1.4	NR	–

	Demographics		Laboratory tests		Biopsy findings	
Patient number	Gender	Age	Creatinine clearance (ml min ⁻¹)	Plasma creatinine (mg dl ⁻¹)	Biopsy diagnosis	Banff scores
Patient #010	M	23	77	1.3	NR	–
Patient #011	M	36	13	5.3	NR	–
Patient #012	M	18	70	1	NR	–
Patient #013	M	27	49	2.4	NR	–
Patient #014	F	46	30	2.8	NR	–
Patient #015	M	20	70	1.3	NR	–
Patient #016	F	20	76	1	NR	–
Patient #017	F	30	50	1.5	EAR	(t-1, i-1, v-0, ptc-2, C4d-0)
Patient #018	F	18	75	1	EAR	(t-1, i-2, v-0, C4d-0)
Patient #019	M	36	58	1.4	EAR	(t-1, i-1, v-0)
Patient #020	F	18	35	2.1	EAR	(t-1, i-1, v-0, ptc-2)
Patient #021	F	20	65	1.1	EAR	(t-2, i-3, v-0, ptc-2)
Patient #022	M	12	80	1.1	EAR	(t-1, i-0, v-0, C4d-0)
Patient #023	M	17	65	1.4	EAR	(t-1, i-1, v-0, C4d-0)
Patient #024	M	12	50	1.7	EAR	(t-1, i-1, v-0, C4d-0)
Patient #025	M	47	27	3.1	EAR	(t-1, i-1, v-0, C4d-0)
Patient #026	M	24	60	1.7	EAR	(t-1, i-1, v-0, C4d-0)
Patient #027	M	35	90	1.4	EAR	(t-1, i-0, v-0)
Patient #028	M	34	55	1.8	EAR	(t-1, i-3, v-0, ptc-1, C4d-0)
Patient #029	M	26	75	1.4	EAR	(t-1, i-1, v-0, ptc-2, C4d-0)
Patient #030	M	36	55	1.7	EAR	(t-3, i-1, v-0, C4d-0)
Patient #031	M	14	80	1.1	EAR	(t-1, i-2, v-0, C4d-0)
Patient #032	M	22	70	1.5	EAR	(t-1, i-0, v-0, C4d-0)
Patient #033	M	15	65	1.4	EAR	(t-1, i-1, v-0, C4d-0)
Patient #034	M	33	55	1.7	EAR	(t-1, i-1, v-0, C4d-0)
Patient #035	F	45	25	2.2	EAR	(t-1, i-1, v-0)
Patient #036	M	17	70	1.5	EAR	(t-1, i-1, v-0)
Patient #037	M	15	75	1.5	EAR	(t-1, i-1, v-0, ptc-2, C4d-0)
Patient #038	M	30	67	1.5	EAR	(t-1, i-0, v-0, C4d-0)
Patient #039	M	26	84	1.1	EAR	(t-1, i-1, v-0, C4d-0).
Patient #040	M	27	86	1.1	EAR	(t-1, i-1, v-0, ptc-0, C4d-0)
Patient #041	M	33	35	2.7	EAR	(t-3, i-2, v-3, ptc-1)
Patient #042	M	20	18	4.6	EAR	(t-1, i-2, v-0, ptc-2)
Patient #043	M	27	65	1.6	EAR	(t-1, i-0, v-0, ptc-2, C4d-0)
Patient #044	M	34	46	2.1	EAR	(t-1, i-1, v-0, ptc-1)
Patient #045	M	27	74	1.2	EAR	(t-1, i-0, v-0, C4d-0)
Patient #046	M	30	45	2.1	EAR	(t-1, i-1, v-0, C4d-0)
Patient #047	M	31	55	1.5	EAR	(t-1, i-1, v-0, C4d-0)
Patient #048	M	26	65	1.5	EAR	(t-1, i-1, v-0, C4d-0)
Patient #049	M	23	60	1.5	EAR	(t-1, i-1, v-0, ptc-3, C4d-0)
Patient #050	M	15	100	0.8	EAR	(t-1, i-0, v-0, C4d-0).

	Demographics		Laboratory tests		Biopsy findings	
Patient number	Gender	Age	Creatinine clearance (ml min ⁻¹)	Plasma creatinine (mg dl ⁻¹)	Biopsy diagnosis	Banff scores
Patient #051	M	54	18	4.7	EAR	(t-1, i-2, v-0, ptc-2, C4d-3)
Patient #052	M	35	63	1.7	EAR	(t-1, i-1, v-0, C4d-0)
Patient #053	M	26	37	2.8	EAR	(t-1, i-0, v-0, ptc-0, C4d-1)

Hint, NR, EAR, t, i, v and ptc denote non-rejection, early AR, tubules, interstitium, blood vessels and peritubular capillaries, respectively.