Transplant International

Transplant International ISSN 0934-0874

ORIGINAL ARTICLE

Assessment of renal allograft fibrosis by acoustic radiation force impulse quantification – a pilot study

Trygve Syversveen,[1] Knut Brabrand,[1] Karsten Midtvedt,[2] Erik H. Strøm,[3] Anders Hartmann,[2,4] Jarl A. Jakobsen[1,4] and Audun E. Berstad[1,4]

1 Department of Radiology, Oslo University Hospital, Rikshospitalet, Oslo, Norway
2 Section of Nephrology, Medical Department, Oslo University Hospital, Oslo, Norway
3 Department of Pathology, Oslo University Hospital, Rikshospitalet, Oslo, Norway
4 Medical Faculty, University of Oslo, Oslo, Norway

Keywords
acoustic radiation force impulse quantification, renal transplant fibrosis, tissue stiffness, ultrasound.

Correspondence
Trygve Syversveen, Department of Radiology, Oslo University Hospital, Rikshospitalet, Sognsvannsveien 20, 0027 Oslo, Norway.
Tel.: +47 23 07 26 00; fax: +47 23 07 26 10;
e-mail: trygve.syversveen@rikshospitalet.no

Received: 24 May 2010
Revision requested: 27 June 2010
Accepted: 9 August 2010
Published online: 2 September 2010

doi:10.1111/j.1432-2277.2010.01165.x

Summary

Chronic allograft nephropathy characterized by interstitial fibrosis and tubular atrophy is a major cause of renal transplant failure. Acoustic radiation force impulse (ARFI) quantification is a promising noninvasive method for assessing tissue stiffness. We evaluated if the method could reveal renal transplant fibrosis. In a prospective study, 30 adult renal transplant recipients were included. ARFI quantification, given as shear wave velocity (SWV), of the renal cortex was performed by two observers. SWV was compared to grade of fibrosis (0–3) in biopsies. The median SWV was 2.8 m/s (range: 1.6–3.6), 2.6 m/s (range: 1.8–3.5) and 2.5 m/s (range: 1.6–3) for grade 0 (n = 12), 1 (n = 10) and grades 2/3 (n = 8) fibrosis respectively. SWV did not differ significantly in transplants without and with fibrosis (grade 0 vs. grade 1, P = 0.53 and grade 0 vs. grades 2/3, P = 0.11). The mean intraobserver coefficient of variation was 22% for observer 1 and 24% for observer 2. Interobserver agreement, expressed as intraclass correlation coefficient was 0.31 (95% CI: −0.03 to 0.60). This study does not support the use of ARFI quantification to assess low-grade fibrosis in renal transplants. ARFI quantification in its present stage of development has also high intra- and interobserver variation in renal transplants.

Introduction

Chronic allograft nephropathy (CAN), now often referred to by the pathological findings interstitial fibrosis (IF) and tubular atrophy (TA), is a major cause of graft loss after kidney transplantation. CAN/IFTA may progress more or less undetected, as sensitive laboratory markers are missing [1,2]. Eventually, there is an increase in serum creatinine and a graft biopsy may then be performed, often too late for meaningful intervention. Increased urinary retinol-binding protein (uRBP) has been reported to be associated with the loss of renal transplant function [3]. uRBP levels were, however, not associated with morphological findings of IF/TA, and whether high levels of uRBP precede structural changes remains to be proven. Means to noninvasively detect and monitor progression of CAN/IFTA at an early stage are thus of great interest. Repeated biopsies are time-consuming, expensive and involve some risk of complication.

Ultrasound (US) elasticity imaging has been applied to various body tissues to evaluate tissue stiffness, and fibrosis tends to increase tissue stiffness [4]. The use of shear wave generation to measure stiffness in renal transplants has been proposed by others [5], and a recent publication reported a significant correlation between renal parenchymal stiffness measured by transient elastography (TE) and the extent of IF [6]. The utility of TE in the detection of liver fibrosis has been reported in several studies, including a study on assessment of liver fibrosis in kidney-transplant patients with chronic viral hepatitis, where an acceptable accuracy for detecting mild liver fibrosis was found [7]. Also, US strain imaging for assessment of tissue stiffness in a renal transplant with mild renal insufficiency and biopsy proven fibrosis showed a threefold

difference in renal cortical strain compared with a normally functioning transplant [8].

One type of elasticity imaging, acoustic radiation force impulse (ARFI) technology, quantitatively assesses shear wave velocity (SWV) of the tissue. Shear waves are created by a short-duration high-intensity acoustic pulse. The classic parameter to describe tissue stiffness is the Young's elastic modulus (Y), which is directly proportional to the square of the SWV [4]. A stiff tissue has a large Y and high SWV. ARFI integrated into conventional US systems is a promising noninvasive method for assessing liver fibrosis as the SWV is accurate, repeatable and strongly correlated with the grade of fibrosis [9,10]. However, we are not aware of any study that has assessed ARFI as a noninvasive tool in detection of renal transplant fibrosis.

The primary aim of this study was to evaluate whether ARFI quantification was able to detect differences in cortical stiffness in renal transplants with and without histologically verified fibrosis. Intra- and interobserver agreement was also analyzed.

Materials and methods

Patients

Patients were prospectively enrolled after obtaining written informed consent. The study was approved by the regional ethics committee.

To include transplants with and without fibrosis, we selected two groups of patients. Group A consisted of patients with low probability of renal fibrosis (normal transplant function assessed from serum creatinine, donor age <50 years, undergoing surveillance protocol biopsy 6 weeks after transplantation). Group B consisted of patients with a higher probability of fibrosis (more than 1 year since transplantation, slow rise in serum creatinine and/or suspected CAN/IFTA). Thirty-seven adult renal transplant recipients referred for US examination including biopsy were eligible for inclusion. Five patients were excluded from analysis because only one observer performed ARFI quantification in these patients, and two patients because of findings of acute rejection in the biopsies. Data from the remaining 30 patients were used in the analysis (group A, $n = 15$, and group B, $n = 15$).

ARFI quantification of renal cortex

All US examinations were performed using an Acuson S2000 US scanner equipped with ARFI quantification (Virtual Touch™ Tissue Quantification package), using a 4-MHz curved linear array transducer. Two investigators performed ARFI measurements (Siemens, Mountain View, CA, USA). The investigators were blinded to the SWV measurements of the other observer, but not to the clinical information. Both had previous training in ARFI quantification. To standardize SWV measurements, skin compression was limited to the weight of the transducer. Each observer performed a total of eight valid measurements from the renal cortex in each patient. With the Virtual Touch™ Tissue Quantification package and 4-MHz transducer, the depth for SWV measurements is limited to 5.5 cm. Therefore, the SWV measurements were obtained from the cortex from an area as close to the skin surface as possible, to assure measurements even from obese patients (Fig. 1). In our study, about half of the patients had the lower pole closest to the skin and measurements were taken from this area, but the exact overlap with the biopsy site cold not be assured. For the rest of the patients, the cortical area closest to the skin was the mid part of the kidney. The distance between the biopsy site and the ARFI measurement site was not more than a few centimeters in any patients in our study. Biopsies from one region of a renal transplant have been found to be representative of estimates of interstitial tissue in explanted human kidney grafts [11]. Hence, lack of overlap between regions for biopsy and SWV measurement was not considered to influence the results. Any acquisition giving a measurement of SWV was considered valid. For technical reasons, some acquisitions did not give an SWV measurement. The ARFI quantification technology and the reasons for such failures have previously been described by others [10]. Success rates were calculated as the ratio between the numbers of validated measurements to the total number of acquisitions. The depth of the region of interest (ROI) for the measurements was recorded.

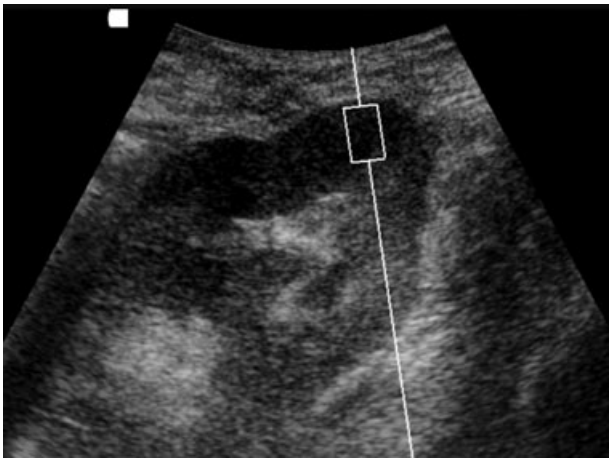

Figure 1 Measurement of share-wave velocity in renal transplant. Rectangle of fixed size (10 × 6 mm) indicates region of interest in the cortex as close to the skin as possible, in this case, in the cortex of the lower pole.

Renal transplant histology and quantification of fibrosis

On the same day as the SWV measurements were performed, US-guided 18-G true-cut biopsies were taken from the peripheral cortex of the lower pole of the transplant. Formalin-fixed, paraffin-embedded tissue was stained with hematoxylin–eosin–safran. The sections were examined by light microscopy by one experienced renal pathologist, blinded to the results of the ARFI measurements. Fibrosis was graded on a 4-point scale (0–3), according to the Banff-scheme [12].

Statistical analysis

For each patient, the median of the total 16 measurements was used as representative for the SWV. The median SWV of transplants without and with fibrosis was compared with Mann–Whitney test. The same comparison was also made for each of the two observers separately using the median of their eight measurements as representative for SWV in each patient. A two-tailed *P*-value of <0.05 was considered statistically significant.

The coefficient of variation (CV = SD/mean × 100%) for repeat SWV measurements of each observer was calculated for each patient. The mean CV value was calculated for each observer. Interobserver agreement was given as the intraclass correlation coefficient [$ICC_{(2.1)}$] [13]. ICC values range from +1 (100% agreement) to −1 (100% disagreement). Interobserver variability was given as a Bland–Altman plot [14].

Results

Table 1 summarizes the patient's demographic and laboratory characteristics. All biopsy specimens were sufficient for histological evaluation. In group A (protocol biopsies), 11 had fibrosis grade 0 and four fibrosis grade 1. In group B, one had fibrosis grade 0, six grade 1, seven grade 2 and one fibrosis grade 3. Based on the total 16 measurements of both observers for each patient, the median SWV was 2.8 m/s (range: 1.6–3.6), 2.6 m/s (range: 1.8–3.5), 2.5 m/s (range: 1.6–3.0) and 1.8 m/s for grade 0, 1, 2 and grade 3 fibrosis respectively. As only one transplant had fibrosis grade 3, grades 2 and 3 were pooled. No significant difference in SWV between transplants without and with fibrosis was detected ($P = 0.53$ and $P = 0.11$ for comparison of fibrosis grade 0 vs. grade 1 and fibrosis grade 0 vs. grades 2/3 fibrosis respectively).

Box and whisker plots for the SWV versus grade of fibrosis for both and each of the observers are shown in Fig. 2 (Fig. 2a both observers, Fig. 2b observer 1 and Fig. 2c observer 2). For observer 1, the SWV was significantly lower for fibrosis grade 2/3 compared with fibrosis grade 0 ($P = 0.02$). No such significant difference was found for observer 2.

Success rate for SWV measurements was 0.90 (range 0.62–1.00). The mean depth of the ROI from the skin surface was 2.7 cm (range: 1.3–4.0 cm). For each of the investigators, the elasticity measurements showed large variation. The mean CV of observer 1 and observer 2 was 22% (range: 7–43%) and 24% (range: 7–40%), respectively. Interobserver agreement, expressed as ICC was 0.31 (95% CI: −0.03 to 0.60). A Bland–Altman plot for the differences in ratings is shown in Fig. 3.

Discussion

We found no significant difference in median SWV between patients without and with renal allograft fibrosis. In contrast to our findings, Arndt *et al.* [6] recently reported a significant correlation between parenchymal stiffness measured by TE and the extent of IF in 20 renal transplants. As both ARFI and TE technology estimate tissue stiffness by tracking of shear wave propagation through the tissue, similar results should be expected for ARFI. There are, however, differences between the two technologies as TE uses vibrations to generate shear waves, whereas ARFI uses short-duration high-intensity

Table 1. Main demographic characteristics and laboratory data of the 30 patients.

	Group A	Group B
Males/females	12/3	11/4
Age, median (range)	47 (21–78) years	62 (40–74) years
Age of donor at transplantation, median (range)	43 (35–51) years*	51 (27–72) years
Time since transplantation at ultrasound examination, median (range)	43 (37–46) days	9.6 (1–17) years
Creatinine in μmol/l median (range)	114 (65–134)	164 (100–259)
Estimated GFR (MDRD) in ml/min/1.73 m^2, median (range)	59 (51–106)	36 (21–64)
Immunosuppressive therapy: calcineurin inhibitor (CNI) and/or mammalian target of rapamycin inhibitor (mTOR)	CNI:14, mTOR:1, both:0	CNI:11, mTOR:3, both:1
Living donor/deceased donor	9/6	7/8

*One patient in group A had an age of 51 years at the time of examination because of a mistake in the inclusion process.

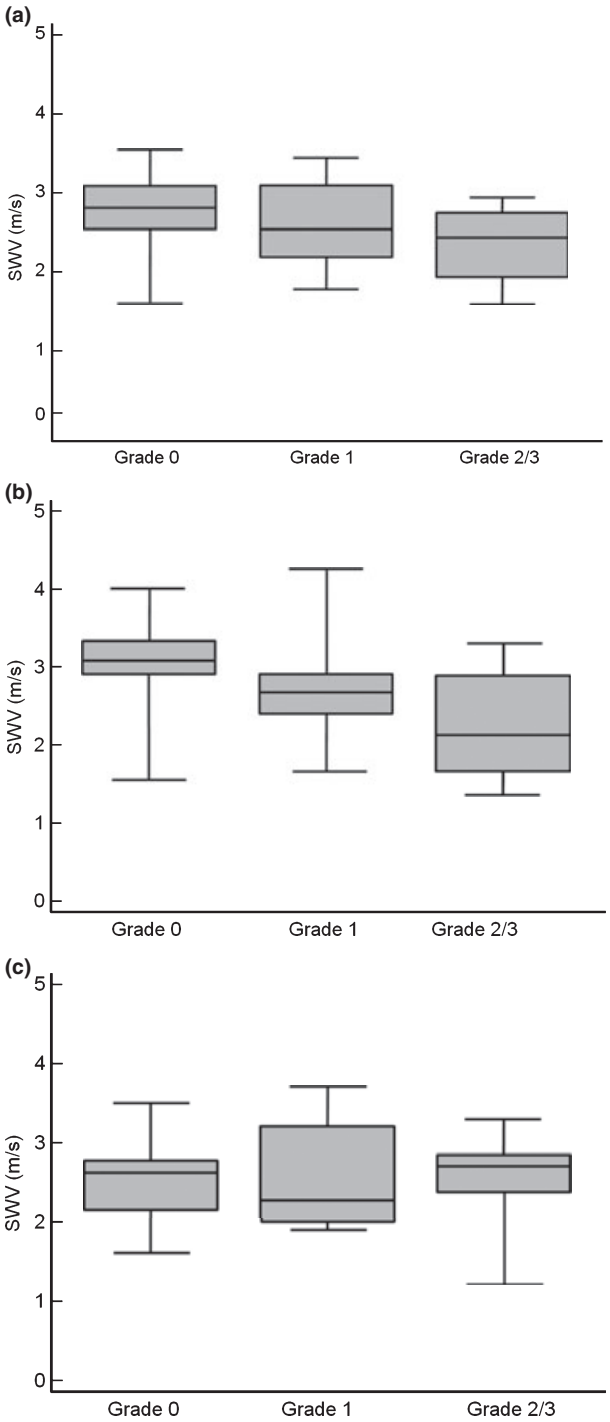

Figure 2 (a) Box- and whiskers plot of shear wave velocities (SWV) for both observers of grade 0 (n = 12), grade 1 (n = 10) and grade 2 and 3 fibrosis (n = 8). The central box represents the values from the lower to upper quartile (25–75 percentile). The middle line represents the median. A line extends from the minimum to the maximum value. (b) Box- and whiskers plot of SWV for observer 1 of grade 0, grade 1 and grade 2 and 3 fibrosis. (c) Box- and whiskers plot of SWV for observer 2 of grade 0, grade 1 and grade 2 and 3 fibrosis.

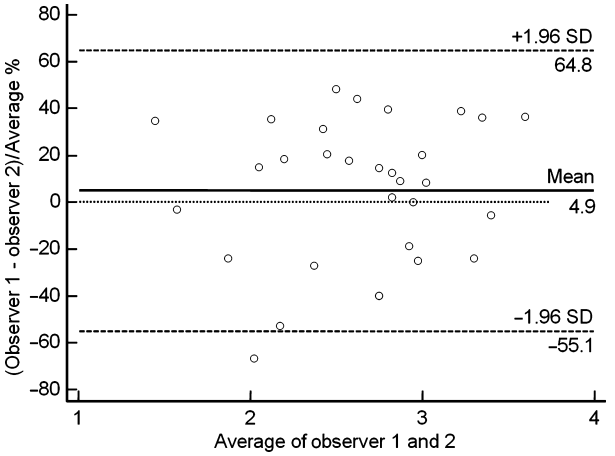

Figure 3 Bland–Altman plot compares independent measurements of shear wave velocity (SWV) of two observers of 30 renal transplants. The SWV of each observer is based on median of eight repeat measurements. The difference of the observers is expressed as percentage deviation from the average of both observers. Horizontal lines are drawn at the mean difference, and at the mean difference plus and minus 1.96 times the standard deviation of the differences.

acoustic pulses. For severe fibrosis of the liver, the diagnostic accuracy of ARFI is comparable to TE, but for assessment of earlier stages of fibrosis, TE performs better than ARFI [10]. Thus, the different results between their study and our study may partly be as a result of different technologies. Also, in the study by Arndt *et al.*, measurements of stiffness were conducted by one observer only, and it was not specified if any of the biopsies showed signs of acute rejection.

In our study, both biopsy and ARFI quantification were performed on the same day. We excluded patients with evidence of acute rejection, as edema and tissue reaction associated with acute rejection might affect the parenchymal stiffness. In the liver, TE is unreliable for detection of cirrhosis in patients with acute liver damage [15].

For observer 1, there was a significant lower SWV in transplants with grades 2/3 fibrosis compared with transplants with grade 0 fibrosis. The reason for this remains unclear as fibrosis in general tends to increase tissue stiffness. However, the same trend was not seen for observer 2, and we therefore assume that the difference between the two observers probably reflects the degree of interobserver variation more than differences in fibrosis.

Indeed, there are concerns about the observer variation in our pilot study. First, the SWV measurements showed large variability for each observer despite the fact that repeat measurements were obtained from the same cortical area. We also found rather low interobserver agreement with an ICC of 0.31. In comparison, TE of the liver has shown excellent interobserver agreement, with ICC as high as 0.98 (95% CI: 0.977–0.987) [16]. The reasons for

the large CV and rather low ICC in our study are not known. It is possible that the tissue stiffness inside the ROI in the renal cortex could be more inhomogeneous than in the liver, as a portion of medulla might have been included in the fixed-size ROI despite our attempts to place it solely in the cortex. Another possible explanation could be that the degree of compression by the transducer differed at repeat measurements and between observers. Elasticity measurements in the liver are usually performed by intercostal scanning and, in this situation, the pressure from the transducer is probably not transferred to the liver surface.

There are several limitations to our study. First, the material is small. This is reflected in the wide CI for ICC (−0.03 to 0.60). A larger material would have given a more accurate estimation of the ICC, but even the upper limit of our CI for the ICC is well below the ICC for TE in the liver. Secondly, our material included only one transplant with fibrosis grade 3, and although that transplant did not differ in SWV, it cannot be excluded that such a high degree of fibrosis can be detected by ARFI measurements. Clinically, however, the detection of low-grade fibrosis is most relevant. Transplants with higher degrees of fibrosis would be expected to have even worse function. At that stage, the transplant often is at a point-of-no return, i.e. it is too late to modify immunosuppression to improve graft outcome [17]. In addition, changes in renal transplants other than fibrosis may have impact on cortical stiffness.

Our study was designed to measure SWV in the renal cortex at a single time in each subject. It remains to be proven whether this technique can detect an increase in fibrosis over time in a specific individual. The overlap between SWV in different grades of fibrosis was, however, large, and the difference in median SWV between the two observers was actually more than 60% (1.4 m/s) in one of the patients. The coefficient of variation in one observer's series of measurements was up to 40%. Hence, an increase in SWV because of progression of fibrosis over time would have to be substantial if it were to be reliably detected by the currently available ARFI quantification. Thus, it does not seem likely that this method would be able to identify individuals with progression of fibrosis over time.

A low reproducibility of the reference (fibrosis grade in graft biopsies) is of course also a potential source of failure to prove the relationship between fibrosis and tissue stiffness. However, despite the limitations of pathology assessment as the 'gold standard', no better alternative exists in the detection of CAN/IFTA. In our study, one and the same pathologist evaluated all the biopsies, and this would at least eliminate inter-observer variation of the reference as a source of error.

Conclusion

Acoustic radiation force impulse quantification was not found to detect differences in cortical stiffness between renal transplants with and without low-grade fibrosis, and at present, renal allograft biopsy remains the gold standard for quantification of fibrosis.

Acoustic radiation force impulse quantification in its present stage of development has also low intra- and inter-observer agreement in renal transplants.

Authorship

TS: designed and performed research/study, collected data, analyzed data, wrote paper. KB: designed and performed research/study, collected data, wrote paper. KM: included patients, collected data, wrote paper. AH: designed research/study, wrote paper. EHS: collected data, wrote paper. JAJ: designed research/study. AEB: analyzed data, wrote paper.

Funding

None.

References

1. Bosmans JL, Ysebaert DK, Verpooten GA. Chronic allograft nephropathy: what have we learned from protocol biopsies? *Transplantation* 2008; **85**(Suppl. 7): S38.
2. Hunsicker LG, Bennett LE. Design of trials of methods to reduce late renal allograft loss: the price of success. *Kidney Int Suppl* 1995; **52**: S120.
3. de Matos AC, Camara NO, de Oliveira AF, *et al.* Functional and morphologic evaluation of kidney proximal tubuli and correlation with renal allograft prognosis. *Transpl Int* 2010; **23**: 493.
4. Garra BS. Imaging and estimation of tissue elasticity by ultrasound. *Ultrasound Q* 2007; **23**: 255.
5. Cosgrove DO, Chan KE. Renal transplants: what ultrasound can and cannot do. *Ultrasound Q* 2008; **24**: 77.
6. Arndt R, Schmidt S, Loddenkemper C, *et al.* Noninvasive evaluation of renal allograft fibrosis by transient elastography – a pilot study. *Transpl Int* 2010; **23**: 871.
7. Alric L, Kamar N, Bonnet D, *et al.* Comparison of liver stiffness, fibrotest and liver biopsy for assessment of liver fibrosis in kidney-transplant patients with chronic viral hepatitis. *Transpl Int* 2009; **22**: 568.
8. Weitzel WF, Kim K, Rubin JM, *et al.* Feasibility of applying ultrasound strain imaging to detect renal transplant chronic allograft nephropathy. *Kidney Int* 2004; **65**: 733.
9. Palmeri ML, Wang MH, Dahl JJ, Frinkley KD, Nightingale KR. Quantifying hepatic shear modulus *in vivo* using acoustic radiation force. *Ultrasound Med Biol* 2008; **34**: 546.

10. Lupsor M, Badea R, Stefanescu H, *et al.* Performance of a new elastographic method (ARFI technology) compared to unidimensional transient elastography in the noninvasive assessment of chronic hepatitis C. Preliminary results. *J Gastrointestin Liver Dis* 2009; **18**: 303.
11. Ellingsen AR, Nyengaard JR, Osterby R, Jorgensen KA, Petersen SE, Marcussen N. Measurements of cortical interstitium in biopsies from human kidney grafts: how representative and how reproducible? *Nephrol Dial Transplant* 2002; **17**: 788.
12. Racusen LC, Solez K, Colvin RB, *et al.* The Banff 97 working classification of renal allograft pathology. *Kidney Int* 1999; **55**: 713.
13. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull* 1979; **86**: 420.
14. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986; **1**: 307.
15. Sagir A, Erhardt A, Schmitt M, Haussinger D. Transient elastography is unreliable for detection of cirrhosis in patients with acute liver damage. *Hepatology* 2008; **47**: 592.
16. Fraquelli M, Rigamonti C, Casazza G, *et al.* Reproducibility of transient elastography in the evaluation of liver fibrosis in patients with chronic liver disease. *Gut* 2007; **56**: 968.
17. Schena FP, Pascoe MD, Alberu J, *et al.* Conversion from calcineurin inhibitors to sirolimus maintenance therapy in renal allograft recipients: 24-month efficacy and safety results from the CONVERT trial. *Transplantation* 2009; **87**: 233.