Kidney International, Vol. 67 (2005), pp. 1515–1520

Correlation of renal histopathology with sonographic findings

Sammy Moghazi, Edria Jones, Jill Schroepple, Kraisith Arya, William McClellan, Randolph A. Hennigar, and W. Charles O'Neill

Renal Division, Department of Medicine; Department of Pathology; and the Rollins School of Public Health, Emory University, Atlanta, Georgia

Correlation of renal histopathology with sonographic findings.

Background. Judgments about irreversible renal disease are frequently based on the sonographic appearance of the kidneys. However, the sensitivity and specificity of sonography in identifying chronic, irreversible disease have never been determined, and the specific pathologic changes that increase renal cortical echogenicity have not been defined.

Methods. We retrospectively compared sonographic parameters (length, quantitative echogenicity, cortical thickness, and parenchymal thickness) to biopsy findings of glomerular sclerosis, tubular atrophy, interstitial fibrosis, and interstitial inflammation in 207 patients.

Results. Echogenicity showed the strongest correlation with all 4 histologic parameters ($r = 0.28$–0.35). Renal size was significantly correlated with glomerular sclerosis ($r = -0.26$) and tubular atrophy ($r = 0.20$). Parenchymal thickness, but not cortical thickness, correlated with tubular atrophy ($r = -0.23$). By multivariate analysis, tubular atrophy and interstitial inflammation, but not interstitial fibrosis, were significant determinants of cortical echogenicity. Severe chronic disease (>50% sclerosed glomeruli or a score of 3 out of 5 or greater for tubular atrophy or interstitial fibrosis) was present in 69% and 47% of patients with combined renal length <20 cm and >20 cm, respectively ($P = <0.05$). For cortical echogenicity >1.0 (>liver echogenicity) and ≤1.0, the proportions of severe disease were 66% and 30%, respectively ($P < 0.001$). Severe disease was present in 86% of patients with combined renal length <20 cm and cortical echogenicity >1.0.

Conclusion. Cortical echogenicity is the sonographic parameter that correlates best with renal histopathology. Although size or echogenicity alone are poor predictors of chronic irreversible disease, the likelihood of treatable disease in small kidneys with increased cortical echogenicity is very low.

Sonography of the kidneys is frequently employed during the evaluation of renal failure. In addition to visualizing a dilated collecting system, sonography provides information on renal size and the thickness and echogenicity of the renal cortex. These parameters are thought to provide an indication of the presence or absence of irreversible parenchymal damage [1–4], and often figure into decisions about whether to perform a renal biopsy, but clinically useful thresholds have never been established. Thinning of the renal cortex and reduced renal size are likely the result of tubular atrophy. Unfortunately, there are very limited normative data on which to judge cortical thinning and, although such data do exist for kidney size, the extent to which a reduction in size correlates with parenchymal disease is unknown. Even less information is available on echogenicity of the renal cortex, which is the backscatter of sound that, in normal cortex, is produced by structures such as glomeruli, vessels, and tubules [5]. It is assumed that collagen present in interstitial fibrosis and glomerulosclerosis is responsible for increased echogenicity [4], but this has never been established. Interstitial inflammation may also increase echogenicity [1]. Furthermore, echogenicity is usually assessed qualitatively by the human eye, which is very unreliable [6, 7], and there is no established normal range. We recently showed that echogenicity of the renal cortex can be reliably quantitated, and established a normal range in a small group of adults [8].

In a previous study comparing sonographic and pathologic findings [4], significant correlations were found between length or cortical echogenicity and glomerular sclerosis or tubular atrophy. Surprisingly, there was no correlation between echogenicity and interstitial fibrosis. However, echogenicity was assessed qualitatively, and no clinically useful thresholds for sonographic parameters were presented. Furthermore, there have been considerable advances in sonographic imaging since this report. Because the extent to which sonographic parameters can predict parenchymal disease has never been established, we correlated histologic findings with sonographic parameters in a large number of patients in whom both sonography and renal biopsy were performed. We sought to validate the correlation between sonographic parameters and parenchymal disease, and to establish thresholds for size and echogenicity that could be useful in making clinical decisions.

Key words: ultrasonography, kidney, renal biopsy, pathology, fibrosis, tubular atrophy, glomerular sclerosis.

Received for publication August 25, 2004
and in revised form October 6, 2004
Accepted for publication October 27, 2004

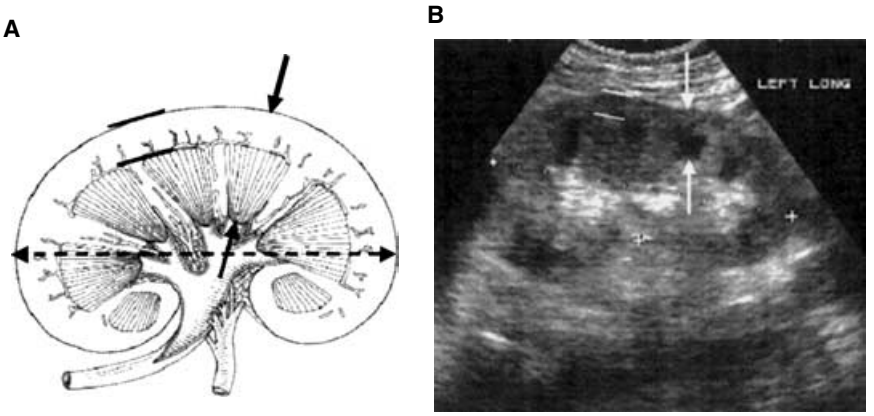

Fig. 1. Sonographic measurements. Diagram of a normal kidney in coronal section (*A*) and longitudinal sonogram of left kidney (*B*) showing the different measurements used. Parallel lines indicate cortical thickness, solid arrows indicate parenchymal thickness, and dashed arrows or dotted line indicate maximal kidney length.

METHODS

Patients

Patients were identified from a database of all sonographic studies performed by the Renal Division since August 1996. Only patients who underwent ultrasound-guided percutaneous renal biopsy of a native kidney, and who had undergone diagnostic sonography of their kidneys between 6 months before and 3 weeks after the biopsy, were included in this study. Of the 419 biopsies performed between August 1996 and September 2002, 207 met the inclusion criteria.

Sonography

All sonograms were performed with an RT3200 scanner (GE Medical Systems, Milwaukee, WI, USA) using a 5.0 MHz wide-frequency band, phased array probe, or with an SDU 450 scanner (Shimadzu, Mountain View, CA, USA) using a 3.0 to 5.0 MHz variable frequency, phased array probe. Output was onto thermal printer paper. All studies were performed by nephrology trainees or faculty thoroughly trained in the procedure, and were interpreted by 1 of 2 experienced nephrologists. Kidney length was determined as the maximum longitudinal dimension, measured at the time of the sonogram, while quantitative determinations of parenchymal thickness, cortical thickness, and echogenicity were performed at a later date for the purpose of this study. The investigators performing these measurements were unaware of the pathologic findings. Parenchymal thickness was determined as the shortest distance from the renal sinus fat to the renal capsule, and cortical thickness was measured as the shortest distance from the base of a medullary pyramid to the renal capsule (Fig. 1). When possible (and in most cases), these measurements were made on the kidney that was biopsied. Echogenicity was quantified as previously described [8]. Briefly, longitudinal images of the right kidney, including adjacent liver, were converted into digital files (8-bit, 300 pixels/in) using a ScanMaker III scanner (Microtek Lab, Redondo Beach, CA, USA) with Photoshop 4.0 software (Adobe Systems, San Jose, CA, USA). Echogenicity was measured as the inverse of the ratio of the mean pixel densities of the renal cortex and adjacent liver using ScionImage software (Scion Corp., Frederick, MD, USA). Because comparison to liver is required, echogenicity was analyzed only in the right kidney. Each sonographic parameter was measured by a single investigator to avoid interobserver variability.

Pathology

Tissue for light microscopic examination was prepared in the conventional manner. Briefly, needle biopsies were fixed in formalin, dehydrated in graded alcohols, cleared in xylene, and infiltrated with paraffin using automated procedures. Paraffin-embedded tissues were cut at 3 microns and stained with hematoxlin and eosin (H&E), periodic acid-Schiff (PAS) reagent, Masson's trichrome, and Jones' methenamine silver stain. Each biopsy contained at least 13 glomeruli. Biopsies were evaluated in terms of 4 parameters: (*1*) glomerular obsolescence, (*2*) tubular atrophy, (*3*) interstitial fibrosis, and (*4*) interstitial inflammation. Each parameter was measured in every biopsy using a semiquantitative grading scale ranging from 0 to 5, where a score of 0 signified that 5% or less of the biopsy was affected; a score of 1, between 6% and 20%; a score of 2, between 21% and 40%; a score of 3, between 41% and 60%; a score of 4, between 61% and 80%; and a score of 5, greater than 80%. All cases were scored by one investigator (E.J.), who was extensively trained by an experienced nephropathologist (R.A.H.). For quality assurance, approximately 40% of cases were reviewed by both investigators, with excellent agreement.

Data analysis

Statistical analysis was performed with SAS version 8.01 software (SAS Institute, Inc., Cary, NC, USA). Univariate analysis is presented as Spearman correlation

Table 1. Patient data

Age *years*	Gender	Kidney biopsied	Interval between sonogram and biopsy *days*
45 (15–82)	Male 98 Female 109	Left 191 Right 16	16 (−16 −148)

Ranges are given in parentheses.

Table 2. Pathologic diagnosis

Diagnosis	Number
Proliferative glomerulonephritis	55
Focal segmental glomerulosclerosis	31
IgA nephropathy	21
Membranous glomerulonephritis	18
Nephrosclerosis	13
Diabetic nephropathy	14
Interstitial disease	9
Thrombotic microangiopathy	10
Amyloidosis	4
Advanced, indeterminant disease	13
Normal	2
Other	11

coefficients. For analysis of sonographic parameters with each of the 4 histologic parameters, a Bonferonni correction of 4 was applied to the *P* value. For comparison of the incidence of severe disease in different populations, significance was determined by the chi-square test using the Yates correction.

RESULTS

Information on the study patients is provided in Table 1. The mean interval between the diagnostic sonogram and the biopsy was 12 days, and 38% of the diagnostic sonograms were performed at the time of the biopsy. Three diagnostic studies were performed after the biopsy (8, 11, and 16 days), but not before any therapy had been instituted. Seven studies were performed more than 2 months before the biopsy. All of these patients had indolent disease that was unlikely to have progressed during this interval. It is standard practice at this institution to biopsy the left kidney, but 8% of the biopsies were performed on the right kidney because the left kidney was absent (1 patient), atrophic (2 patients), or contained cysts (1 patient). In the remaining 11 patients, the reason for using the right kidney was not specified. The most common pathologic diagnosis (Table 2) was proliferative glomerulonephritis, of which 36% were due to systemic lupus erythematosis (SLE). SLE was also the etiology in 4 of the patients with membranous nephropathy. In 6%, the renal disease was too advanced to make a specific diagnosis. Some patients had more than one diagnosis and were categorized as having each.

Sonographic data are presented in Table 3. Since renal length varies with body height and those data were unavailable, it was not possible to determine which kidneys were outside the normal range. Mean renal length was in the normal range for subjects aged 40 to 50 years (#180), but 18% of right kidneys and 15% of left kidneys were less than 10 cm. The mean difference in length between the 2 kidneys was 7 mm, due principally to a slightly larger left kidney, which is normal [9]. In only 7 patients (3.4%) was there a size discrepancy >2 cm, suggestive of asymmetric disease [10]. Mean echogenicity was increased compared to our previously established normal range of 0.810 to 0.987, and was above this range in 59% of the patients in whom it could be measured. This was consistent with the original qualitative assessment of echogenicity as being normal (less than the liver) in only 32%. Echogenicity could not be quantified in 39 studies because the images were not available or were not suitable for quantitative analysis. The qualitative assessment of echogenicity in these studies did not differ from that in the other studies.

Cortical thickness could not be determined in 82 of the biopsied kidneys (77 left kidneys), primarily because the medullary pyramids were not sufficiently visible. In 12 of these patients, thickness could be measured in the other kidney, and this value was used. Based on a normal range of 8 to 11.5 mm established in a small study of transplant donors [11], cortical thickness was reduced in 41% and increased in 3% of studies. Parenchymal thickness could be measured on a slightly larger number of kidneys, but had a wider range, probably because of the variability within kidneys and the imprecision of its measurement. A normal range has not been established, but it is probably centered around 15 to 16 mm [12]. There was no significant correlation between any of the sonographic parameters other than between parenchymal and cortical thickness.

Composite pathologic data are presented in Table 4. Each parameter varied through the full range, and there were 7 biopsies with a zero score on each parameter. Not surprisingly, all parameters were strongly associated with each other, with correlation coefficients ranging from 0.41 for glomerulosclerosis and interstitial inflammation to 0.89 for tubular atrophy and interstitial fibrosis. Correlations between sonographic parameters and pathologic findings are shown in Table 5. Echogenicity correlated most strongly with tubular atrophy and interstitial inflammation, with slightly weaker correlations with glomerular sclerosis and interstitial fibrosis. These relationships are plotted in Figure 2. Of the 62 patients with normal renal cortical echogenicity ($\leq$0.987), mean scores for glomerulosclerosis ($15 \pm 2.2\%$), interstitial fibrosis (1.4 ± 0.19), and tubular atrophy (1.1 ± 0.14) were significantly lower than the scores in patients with increased echogenicity ($29 \pm 2.4\%$, 2.6 ± 0.15, and 2.1 ± 0.15). The correlation between pathologic findings and

Table 3. Sonographic findings

	Length *cm*			Cortical echogenicity	Cortical thickness	Parenchymal thickness
	Left	Right	Diff.[a]	Kidney/liver	*mm*	*mm*
Mean	10.9	10.9	0.7	1.04	8.3	17.1
Median	11.0	11.0	0.6	1.02	8.2	16.3
Range	7.5–15.0	7.6–14.5	0–4.6	0.71–1.89	4.7–12.5	7.0–33.7
N	205	207	205	168	137	145

[a]Difference in lengths between kidneys.

Table 4. Pathologic findings

	Sclerosed glomeruli %	Interstitial fibrosis	Tubular atrophy	Interstitial inflammation
Mean	24.6	2.3	1.8	1.1
Median	15.0	2	2	1
Range	0–100	0–5	0–5	0–5

Table 5. Spearman correlation coefficients for the relationships between sonographic parameters and pathologic parameters

	Sclerosed glomeruli	Interstitial fibrosis	Tubular atrophy	Interstitial inflammation
Kidney length	−0.26	−.14	−0.20	0.10
	$P = 0.0002$	$P = 0.0048$	$P = 0.0042$	$P = 0.15$
Echogenicity	0.30	0.35	0.28	0.34
	$P < 0.0001$	$P < 0.0001$	$P = 0.0003$	$P < 0.0001$
Parenchymal thickness	−0.20	−.19	−0.23	−0.14
	$P = 0.014$	$P = 0.025$	$P = 0.0061$	$P = 0.097$
Cortical thickness	−0.15	−0.005	−0.12	0.005
	$P = 0.08$	$P = 0.95$	$P = 0.15$	$P = 0.95$

length was significantly weaker, but was better than for parenchymal thickness. There was no significant correlation between cortical thickness and any histopathologic parameter.

Because the histologic parameters were strongly linked, multivariate analysis was performed to determine their relative contributions to echogenicity (Table 6). In this case, interstitial fibrosis was no longer correlated with echogenicity, and only tubular atrophy and interstitial inflammation showed significant independent contributions. The value B indicates the degree to which each parameter contributed to echogenicity. In the case of tubular atrophy, each unit increase added 0.051 to the echogenicity. That tubular atrophy and not interstitial fibrosis is an independent determinant of echogenicity is illustrated by the 2 cases in which the scores for interstitial fibrosis and tubular atrophy were widely disparate. In a patient with rapidly progressive glomerulonephritis and scores for interstitial fibrosis and atrophy of 0 and 4, respectively, echogenicity was 1.021, while in a patient with chronic thrombotic microangiopathy and scores for interstitial fibrosis and atrophy of 4 and 1, echogenicity was 0.929 (normal). In both cases, the interstitial inflammation score was 1. The composite r^2 for the multivariate analysis was 0.20, indicating that histologic parameters accounted for a fifth of the variation in echogenicity.

Two sonographic thresholds (combined kidney length <20 cm and echogenicity greater than liver) were evaluated to determine their clinical utility (Table 7). For this purpose, patients with scores of at least 51% for glomerular sclerosis, 3 for interstitial fibrosis, or 3 for tubular atrophy were judged to have severe chronic disease, and 105 cases met this definition. The incidence of severe disease was approximately 2/3 in small kidneys or in echogenic kidneys, and was significantly greater than in larger or nonechogenic kidneys. Combining the thresholds identified a smaller number of patients, but in whom the incidence of severe chronic disease was very high.

DISCUSSION

This study demonstrates that sonographic determination of renal length and cortical echogenicity correlates with chronic, irreversible renal disease. For each histologic measure of disease, the correlation was substantially stronger with echogenicity than with renal length. Echogenicity was correlated with each histologic parameter, but only tubular atrophy and interstitial inflammation remained significant in a multivariate analysis. Since it is widely assumed that collagen fibrils contribute importantly to the acoustic backscatter of tissues [2, 5, 13], it was surprising that interstitial fibrosis was not an independent determinant of echogenicity. It is of interest that a previous study using qualitative assessment of echogenicity also found no correlation with interstitial fibrosis [4]. Possible mechanisms by which tubular atrophy could increase renal cortical echogenicity include thickening of the tubular basement membranes or luminal dilatation of the remaining tubules [14]. We have previously shown that echogenicity varies directly with diuresis [8], presumably due to changes in tubular caliber. The pathologic parameters accounted for only 20% of the variation in echogenicity, and it is likely that the remaining 80% is related to variability in the sonographic technique or image analysis. The possibility that the biopsy is not representative of the entire kidney may also contribute to the variability. One limitation is that echogenicity could only be determined for the right kidney, whereas most of the biopsies were obtained from the left kidney.

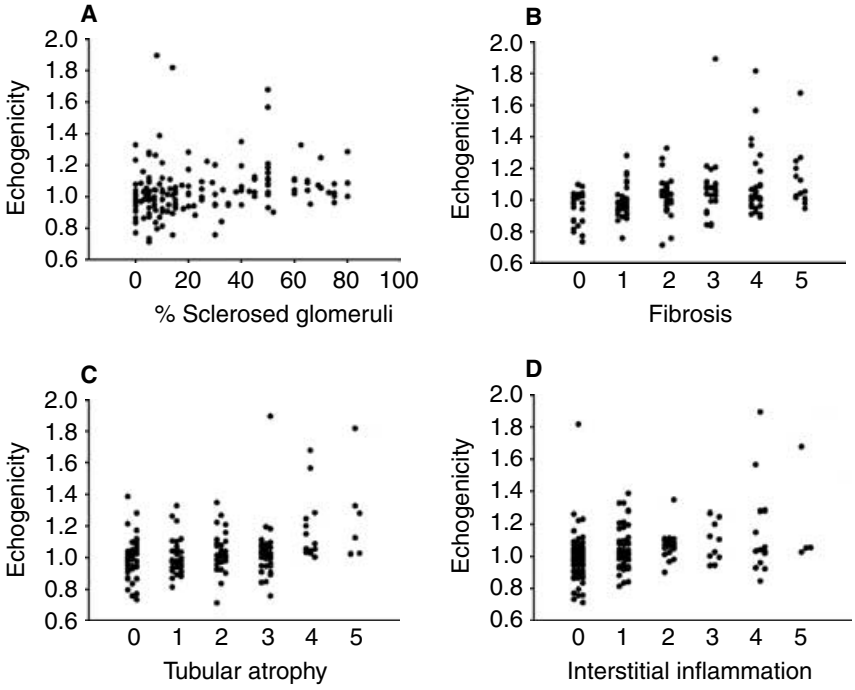

Fig. 2. Relationship between renal cortical echogenicity and pathologic findings.

Table 6. Multivariate analysis with echogenicity as the dependent variable

	Beta	Std. error	*P* value
Tubular atrophy	0.051	0.022	<0.001
Interstitial inflammation	0.040	0.015	0.0073
Glomerular sclerosis	−0.0016	0.0009	0.0556
Interstitial fibrosis	−0.0039	0.021	0.85

However, the disorders present in these patients should affect both kidneys equally.

Neither parenchymal thickness nor cortical thickness provided any better correlation with histopathology than did a simple measurement of renal length. The poor correlations with these direct measures of parenchymal size can probably be explained by the difficulty in accurately and reproducibly measuring either thickness. The diagnostic studies reported here were not performed with the intent of accurately determining parenchymal or cortical thickness so that optimal images for this purpose were often not obtained. A prior study found a weak correlation between parenchymal thickness and histology that was not strong enough to influence the decision to perform a biopsy [15].

Although kidney length and echogenicity were correlated with histologic findings of chronic disease, neither parameter alone was a good discriminator. The specificity for severe chronic disease was only two thirds. Very low echogenicities excluded very severe glomerular sclerosis, interstitial fibrosis, or tubular atrophy, but elevated echogenicities were not useful. The lack of discrimination by increased echogenicity may be related to interstitial inflammation, which was an important cause of increased echogenicity, but is potentially reversible and not necessarily an indication of chronic disease. Combining kidney length and echogenicity provided much better discrimination, with 86% of patients having combined kidney length <20 cm and echogenicity >1.0 (greater than the liver), demonstrating severe chronic disease in the biopsy. Although these conclusions are hampered by the small numbers, it is likely that specificity is even greater in the overall population of renal disease. There were undoubtedly a large number of patients with small echogenic kidneys and severe disease in whom biopsy was not performed, whereas almost all patients with normal-appearing kidneys are likely to have undergone biopsy.

CONCLUSION

This study is the first systematic comparison of sonographic findings with renal histopathology. It shows that renal cortical echogenicity is determined primarily by tubular atrophy and interstitial inflammation, and is the sonographic parameter that correlates best with pathologic findings. While reduced renal length or increased echogenicity alone lack specificity for severe chronic disease, together they demonstrated good specificity, supporting the current clinical practice. Measurement of parenchymal or cortical thickness was not useful.

Table 7. Histopathology at different thresholds for sonographic parameters

Parameter	*N*	% Severe disease	% Sclerosed glomeruli	Interstitial fibrosis	Tubular atrophy
Combined length <20 cm	32	69[a]	36 ± 4	3.1 ± 0.3	2.3 ± 0.3
Combined length ≥20 cm	175	47	23 ± 2	2.1 ± 0.1	1.7 ± 0.1
Echogenicity >1.0	99	66[b]	30 ± 3	2.7 ± 0.2	2.1 ± 0.2
Echogenicity ≤1.0	69	30	14 ± 2	1.6 ± 0.2	1.2 ± 0.1
Combined length <20 cm and echogenicity >1.0	21	86[b]	46 ± 5	3.5 ± 0.2	2.7 ± 0.3
Combined length ≥20 cm or echogenicity ≤1.0	147	46	20 ± 2	2.1 ± 0.1	1.6 ± 0.1

Errors are standard errors.
[a] $P < 0.05$.
[b] $P < 0.001$.

Reprint requests to W. Charles O'Neill, M.D., Emory University School of Medicine, Renal Division WMB 338, 1639 Pierce Dr., Atlanta, GA 30322.
E-mail: woneill@emory.edu

REFERENCES

1. Rosenfield AT, Siegel NJ: Renal parenchymal disease: Histopathologic-sonographic correlation. *AJR* 137:793–798, 1981
2. Rosenfield AT, Taylor KJW, Crade M, *et al*: Anatomy and pathology of the kidney by gray scale ultrasound. *Radiology* 128:737–744, 1978
3. Paivansalo M, Huttunen K, Suramo I: Ultrasonographic findings in renal parenchymal diseases. *Scand J Urol Nephrol* 19:119–123, 1985
4. Hricak H, Cruz C, Romanski R, *et al*: Renal parenchymal disease: Sonographic-histologic correlation. *Radiology* 144:141–147, 1982
5. Insana MF, Hall TJ, Fishback JL: Identifying acoustic scattering sources in normal renal parenchyma from the anisotropy in acoustic properties. *Ultrasound Med Biol* 17:613–626, 1991
6. Lamont AC, Graebe AC, Pelmore JM, *et al*: Ultrasound assessment of renal cortical brightness in infants: Is naked eye evaluation reliable? *Invest Radiol* 25:250–253, 1990
7. Eggert P, Debus F, Kreller-Laugwitz G, *et al*: Densitometric measurement of renal echogenicity in infants and naked eye evaluation: A comparison. *Pediatr Radiol* 21:111–113, 1991
8. Manley JA, O'Neill WC: How echogenic is echogenic? Quantitative acoustics of the renal cortex. *Am J Kidney Dis* 37:706–711, 2001
9. O'Neill WC, Sonography of the normal kidney, in *Atlas of Renal Ultrasonography*, edited by O'Neill WC, Philadelphia, WB Saunders, 2000
10. O'Neill W: Sonographic evaluation of renal failure. *Am J Kidney Dis* 35:1021–1038, 2000
11. Raj DSC, Hoisala R, Somiah S, *et al*: Quantitation of change in the medullary compartment in renal allograft by ultrasound. *J Clin Ultrasound* 25:265–269, 1997
12. Emamian SA, Nielsen MB, Pedersen JF, *et al*: Kidney dimensions at sonography: Correlation with age, sex, and habitus in 665 adult volunteers. *AJR* 160:83–86, 1993
13. Price RR, Jones TB, Goddard J, *et al*: Basic concepts of ultrasonic tissue characterization. *Radiol Clinics North Amer* 18: 21–30, 1980
14. Lajoie G, Silva FG: Approach to the interpretation of renal biopsy, in *Renal Biopsy Interpretation*, edited by Silva FG, D'Agati VD, Nadasdy T, New York, Churchill Livingstone, 1996
15. Roger SD, Beale AM, Cattell WR, *et al*: What is the value of measuring renal parenchymal thickness before renal biopsy? *Clin Radiol* 49:45–49, 1994