US for Detecting Renal Calculi with Nonenhanced CT as a Reference Standard[1]

Keir A. B. Fowler, MD
Julie A. Locken, MD
Joshua H. Duchesne
Michael R. Williamson, MD

Index terms:
Kidney, calculi, 811.811, 813.811
Kidney, CT, 81.12111, 81.12115
Kidney, US, 81.1298

Published online before print
10.1148/radiol.2221010453
Radiology 2002; 222:109–113

Abbreviation:
ESWL = extracorporeal shock-wave lithotripsy

[1] From the Department of Radiology, University of New Mexico School of Medicine, University of New Mexico Health Sciences Center, 2211 Lomas Blvd, Albuquerque, NM 87131. Received February 14, 2001; revision requested March 20; revision received May 11; accepted July 5. **Address correspondence to** K.A.B.F. (e-mail: *keirfowler@home.com*).

PURPOSE: To determine the sensitivity and specificity of ultrasonography (US) for detecting parenchymal and renal pelvis calculi and to establish the accuracy of US for determining the size and number of calculi.

MATERIALS AND METHODS: A total of 123 US and computed tomographic (CT) examinations were compared retrospectively for the presence of renal calculi. The sensitivity of US was determined for individual calculi and at least one calculus per examination. Retrospective findings were compared with the original US interpretation. The sizes of calculi in longest axis were compared on US and CT images, and the US detection of calculi in the left and right kidneys was compared. The use of US for detecting the full extent of calculus burden was evaluated in patients with multiple calculi.

RESULTS: US depicted 24 of 101 calculi identified at CT, yielding a sensitivity of 24% and a specificity of 90%. There was no substantial difference for the detection of calculi in the right and left kidneys. The sensitivity of US for any calculi in a patient was 44%, equal to that of the original US interpretation. US enabled identification of 39% of patients with multiple calculi and demonstrated all calculi in 17% of these patients. The mean size of calculi detected with US was 7.1 mm ± 1.2 (95% CI); 73% of calculi not visualized at US were less than 3.0 mm in size. Calculus size based on US and CT measurements was concordant in 79% of cases and differed by a mean of 1.5 mm ± 0.7.

CONCLUSION: US is of limited value for detecting renal calculi.

Nonenhanced helical computed tomography (CT) has become the primary imaging modality for evaluating acute flank pain and suspected renal stone disease. The high sensitivity (97%) and specificity (96%) of helical CT for depicting genitourinary calculi has been established (1), and CT is of particular value for detecting ureteral calculi, which often are not visualized with other imaging modalities. With the increased use of CT, ultrasonography (US) and intravenous pyelography have begun to play a secondary role in the evaluation of genitourinary calculi.

Nonetheless, US continues to be performed in the setting of acute flank pain or nephrolithiasis for the detection of calculi in the renal pelvis and parenchyma. US is also performed to identify fragmented renal calculi after extracorporeal shock-wave lithotripsy (ESWL). The sensitivity of US for detecting renal calculi has been reported to be as high as 96% compared with that of abdominal radiography and conventional tomography (2). However, the true sensitivity of US for renal calculi may be substantially less given evidence that radiography is less sensitive than previously thought (3).

The sensitivity of US for detection of renal calculi compared with that of helical CT is unclear. Establishing the sensitivity of US for renal calculi will allow informed decisions regarding which type of imaging examination to perform for a given clinical situation. Whereas study authors have evaluated the sensitivity of US for ureteral calculi relative to that of nonenhanced helical CT (4), we are not aware of prior studies in which US and nonenhanced helical CT were compared for sensitivity for calculi within the renal pelvis or renal parenchyma.

The twofold purpose of our study was, with nonenhanced CT as a standard of reference, to perform a retrospective analysis to determine the sensitivity and specificity of US for

Author contributions:
Guarantors of integrity of entire study, K.A.B.F., M.R.W.; study concepts and design, all authors; literature research, K.A.B.F., J.H.D., M.R.W.; clinical studies, all authors; data acquisition, K.A.B.F., J.H.D.; data analysis/interpretation, K.A.B.F., M.R.W.; statistical analysis, K.A.B.F., M.R.W.; manuscript preparation, K.A.B.F., M.R.W.; manuscript definition of intellectual content, K.A.B.F., J.A.L., M.R.W.; manuscript editing, K.A.B.F., M.R.W.; manuscript revision/review, K.A.B.F., J.A.L., M.R.W.; manuscript final version approval, all authors.

detecting parenchymal and renal pelvis calculi and to establish the accuracy of US for determining the size and number of calculi.

MATERIALS AND METHODS

Approval from our institutional review board was obtained for the purposes of this study. Informed patient consent is not required for the review of images at our institution.

Our study included 123 patients (45 male and 78 female patients; mean age, 45 years; age range, 14–88 years) who for clinical reasons underwent both nonenhanced helical CT and US of the kidneys at our institution between December 1997 and July 2000. Patients were selected from those undergoing a standardized CT protocol for suspected renal calculi who also underwent renal US within 30 days preceding or following CT. Patients who had previously undergone renal transplantation were excluded from the study. The mean interval between CT and US examinations was 5.9 days ± 1.5 (95% CI).

Nonenhanced helical CT was performed by using a model PQ 2000 scanner (Picker, Cleveland, Ohio) and a dedicated protocol with 5.0-mm collimation and 1.0 pitch (120–140 kVp, 120–140 mA). Scanning was performed from the upper abdomen through the pubis, with images reconstructed at 5.0-mm intervals. US was performed by using new-generation scanners (model 128, Acuson, Mountain View, Calif; and model 5000, ATL, Bothell, Wash) and consisted of either dedicated renal or abdominal imaging. US included evaluation of the kidneys in multiple anatomic planes.

In this article, focal high-attenuating opacities at CT or shadowing echogenic foci at US are termed as calculus or calculi because CT and US cannot enable reliable distinction of calcium deposition from a concretion of different materials with similar attenuation or echogenicity, respectively. Terms such as *nephrocalcinosis* and *nephrolithiasis* have been avoided because the purpose of this study was to determine the use of US in detecting renal pelvis and parenchymal calculi, regardless of the underlying cause of calculus formation.

US and CT examinations (one per patient) were reviewed in a blinded retrospective manner by two staff radiologists (M.R.W., J.A.L.) who subspecialize in body imaging. The reviewing radiologists worked as a team and provided a consensus interpretation for each image. For each patient, the US images were reviewed prior to the CT examination. The calculi size (longest axis) and number were recorded for the US and CT examinations. The location of each calculus was recorded as being in either the right or the left kidney. No distinction was made between calculi within the renal pelvis and renal parenchyma. Renal calculi were diagnosed on US images on the basis of focal echogenicity with acoustic shadowing in the renal parenchyma or renal pelvis. Punctate high-attenuating foci in the renal parenchyma or renal pelvis were used as criteria for the diagnosis of renal calculi on CT scans. Calculi in the ureter or bladder were not included in this study.

The sensitivity of US for calculi in the renal parenchyma or pelvis was calculated by using CT as a reference. The sensitivity for the presence or absence of any renal calculi for each patient was also calculated for comparison with the original US interpretations at the time the studies were performed. Calculi were classified according to size in groups of 0.0–3.0 mm, 3.1–7.0 mm, and greater than 7.0 mm, since it has been shown that calculi size affects patient therapy (5). Calculi were termed as concordant if both CT and US size measurements were within the same group. A calculus was considered discordant if the size group based on US was different from that determined with CT measurement. The use of US for detecting all calculi in patients with multiple calculi was evaluated, and the mean number of calculi missed in this subset of patients was determined. The sensitivity of US for calculi in the left and right kidneys was evaluated to determine if US is less effective for detecting left-sided calculi, since complete depiction of the left kidney is often difficult because of its more superior position requiring the use of an intercostal window.

After retrospective review of all images, the initial interpretation of the US images was performed by reviewing (K.A.B.F.) the original radiology report for each US examination. The findings in the reports were classified as positive for renal calculi, negative for renal calculi, or indeterminate. The sensitivity and specificity for any renal calculi were determined by using the retrospective CT scan interpretation as a reference. Cases in which the report was indeterminate for the presence of renal calculi were excluded from these calculations. The sensitivity for individual calculi was not determined because in many cases the original US report lacked sufficient information regarding calculus location or size to allow reliable correlation with the retrospective CT interpretation.

Mean values for calculi size were calculated by using 95% CIs. For measures of sensitivity, specificity, positive and negative predictive values, and accuracy, a 95% CI was calculated by using a normal approximation of the binomial distribution (6).

TABLE 1
Detection of 101 Renal Calculi

	US Findings	
CT Findings	Positive	Negative
Positive	24	77
Negative	9	78

Note.—The sensitivity of US was 24% (95% CI: 18%, 30%), and the specificity was 90% (95% CI: 85%, 94%).

TABLE 2
Detection of Any Renal Calculi in 39 Patients with Calculi at CT

	US Findings	
CT Findings	Positive	Negative
Positive	17	22
Negative	6	78

Note.—The sensitivity of US was 44% (95% CI: 35%, 53%), and the specificity was 93% (95% CI: 88%, 98%).

RESULTS

Of the 123 patients, 39 (32%) had 101 renal calculi identified on nonenhanced helical CT scans. The patients with renal calculi included 26 female patients and 13 male patients, with a mean age of 42 years (age range, 14–82 years). The mean calculus size (longest axis) was 4.2 mm ± 0.4, with a range of 0.5–26.0 mm. Of these 39 patients, 18 (45%) had multiple renal calculi on CT scans, with a mean of 4.4 calculi ± 1.4 per patient and a mean size of 3.0 mm ± 0.7 (size range, 0.5–15.0 mm).

As shown in Table 1, US demonstrated 24 of the 101 renal calculi identified on CT images, corresponding to a sensitivity of 24% (95% CI: 18%, 30%) and a specificity of 90% (95% CI: 85%, 94%). The accuracy of US for the detection of individual renal calculi was 54% (95% CI: 47%, 61%), with a positive predictive

TABLE 3
US Detection of 101 Renal Calculi according to Size

Calculi Size (mm)	Total No. of Calculi	Detected	Missed	Sensitivity*
0–3.0	64	8	56	13
3.1–7.0	23	6	17	26
>7.0	14	10	4	71

* Sensitivities are percentages.

TABLE 4
Calculus Size Agreement between CT and US in 24 Cases

	Calculus Size at US (mm)		
Calculus Size at CT (mm)	0–3.0	3.1–7.0	>7.0
0–3.0	6	2	0
3.1–7.0	2	3	1
>7.0	0	0	10

Note.—Calculus size was concordant in 19 (79%) of the 24 cases.

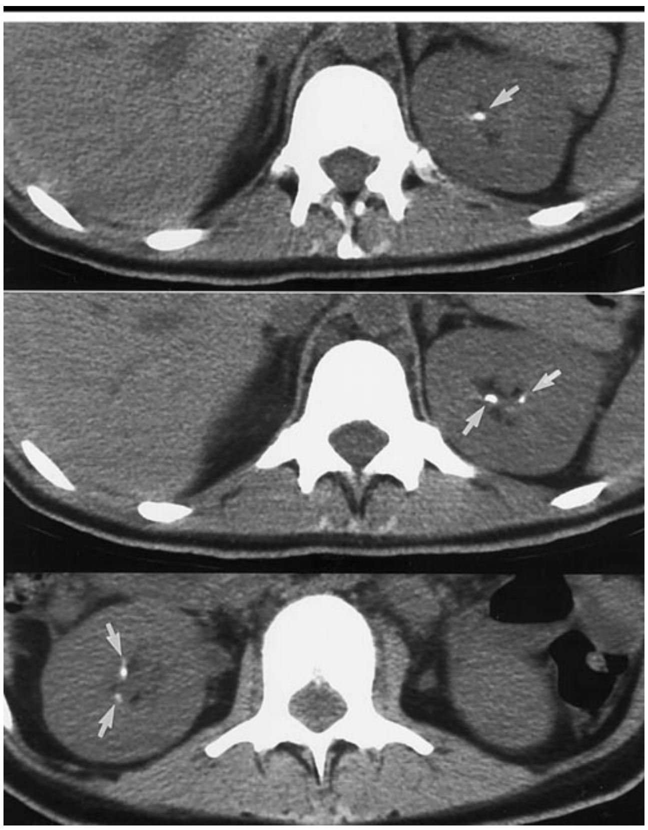


Figure 1. Multiple transverse nonenhanced CT images demonstrate numerous 3–5-mm renal calculi (arrows) that were not depicted at US.

value of 73% (95% CI: 67%, 79%) and a negative predictive value of 50% (95% CI: 43%, 57%). The mean size of calculi detected with US was 7.1 mm ± 1.2.

In the 39 patients with calculi on the basis of CT findings, US depicted at least one calculus each in 17 patients (Table 2). The sensitivity and specificity of US on an examination-by-examination basis for any renal calculi were 44% (95% CI: 35%, 53%) and 93% (95% CI: 88%, 98%), respectively. The accuracy of US for detecting any renal calculi in a given patient was 77% (95% CI: 70%, 84%).

US depicted seven of 34 calculi in the right kidney, corresponding to a sensitivity of 21% (95% CI: 7%, 35%). Seventeen of 67 calculi in the left kidney were detected with US, yielding a sensitivity of 25% (95% CI: 15%, 35%).

US enabled identification of seven (39%) of the 18 patients with multiple calculi. This subset of patients had a total of 80 calculi, 14 (18%) of which were identified at US. In the 18 patients with multiple calculi on CT scans, US demonstrated all calculi in three (17%) patients. With US, a mean of 3.7 ± 1.7 calculi were missed per patient in this group of patients; the mean size of the missed calculi was 3.4 mm ± 0.7 (range, 0.5–15.0 mm).

Of the 101 calculi identified on CT scans, 77 were not visualized at US (Fig 1). Table 3 demonstrates that a majority (56 [73%] of 77) of missed calculi were 3.0 mm or less in size. Slightly fewer than half (16 [43%] of 37) of the calculi larger than 3.0 mm were depicted at US. The mean size of calculi missed with US was 3.3 mm ± 0.6, substantially smaller than the mean size of detected calculi. The sensitivity for small calculi (0.0–3.0 mm) was 13% (95% CI: 5%, 21%). The sensitivities for medium (3.1–7.0-mm) and large (>7.0 mm) calculi were 26% (95% CI: 8%, 44%) and 71% (95% CI: 47%, 95%), respectively.

As shown in Table 4, calculus size based on US and CT images was concordant in 19 (79%) of 24 calculi. With US measurement, three calculi were incorrectly grouped in a larger size group, including a 5-mm calculus that measured 8 mm on US images. Two calculi were incorrectly classified in a smaller size group on the basis of US measurements. The mean difference in the sizes of calculi as determined at CT and US was 1.5 mm ± 0.7.

On US images, nine echogenic foci were present that were originally interpreted as renal calculi but did not correlate with calculi on CT scans. In one case, a 15-mm echogenic focus depicted with US was found on CT images to be a ureteral stent. Renal vascular calcifications were misinterpreted as renal calculi on US images in two patients (Fig 2). In the remaining cases, the cause of the echogenic foci was not clear.

The results of the 123 original US reports reviewed are shown in Table 5. Seven examinations were indeterminate for renal calculi and were excluded from further evaluation. Based on the remaining 116 original reports, the sensitivity and specificity of US for any renal calculi were 44% (95% CI: 35%, 53%) and 98%

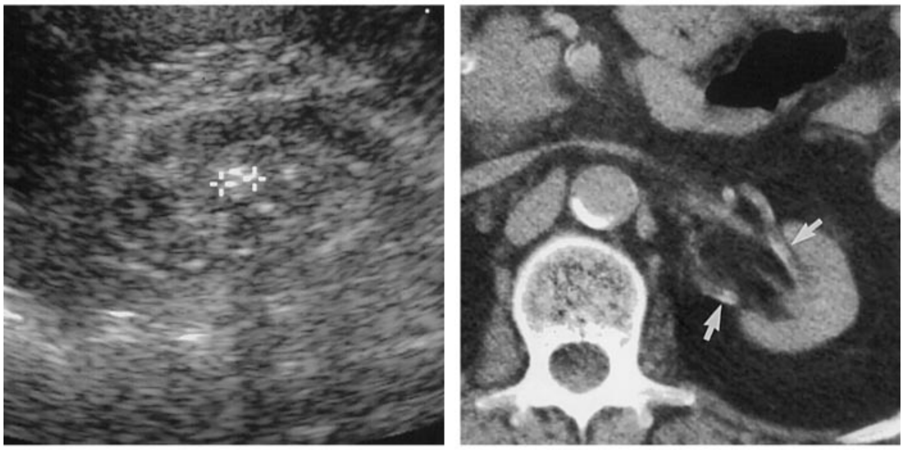
(95% CI: 95%, 100%), respectively. The positive predictive value was 88% (95% CI: 82%, 94%), and the negative predictive value was 81% (95% CI: 74%, 88%).

DISCUSSION

The high sensitivity of nonenhanced helical CT for genitourinary calculi has been established (7), and this modality is viewed by many to be preferred for depicting renal colic and evaluating renal stone disease (8). Nonenhanced CT enjoys clear advantages for evaluation of ureteral calculi that are often difficult to visualize with US or radiography because of overlying bowel gas and adjacent bone structures. Yilmaz and colleagues (4) have demonstrated the superiority of CT for the detection of ureteral calculi compared with both US and intravenous urography. In their study, the sensitivity of US for ureteral calculi was found to be 19% compared with 94% for nonenhanced CT.

Although CT is generally acknowledged as superior for evaluation of ureteral calculi, the difference between US and CT for evaluation of renal pelvis and parenchymal calculi has not, to our knowledge, been established. Previous studies (2,9) in which radiography and conventional tomography are compared with US indicated the sensitivity of US for renal calculi to be as high as 99%. These initial study findings may have been misleading, since Levine et al (3) have shown that many calculi detected at nonenhanced helical CT are missed at radiography. Beyond the evaluation of renal stone disease, the sensitivity of US for renal calculi in general is important because US images are often obtained for other renal disorders in which calcium deposits occur in the form of calculi.

Our data indicate that US is of limited value for the detection of renal calculi. Of the 101 renal pelvis or parenchymal calculi identified on CT scans, only 24 (24%) were depicted on renal US images. No substantial difference in sensitivity was observed between the right and left kidneys, consistent with the prior findings of Middleton et al (2). The sensitivity of US for renal calculi in the current study is substantially lower than that in prior studies (2,9) in which US was compared with radiography and conventional tomography. This finding suggests that with both radiography and CT, a substantial number of renal calculi are missed that are easily detected with CT. Data from a previous study by Sommer et al (10) in which ureteral calculi were evaluated suggest similar difficulties identifying renal calculi at US. In that study, seven renal calculi were identified at CT, whereas a single renal calculus was detected at US in the same group of patients.



a. b.

Figure 2. Vascular calcifications in this patient were misinterpreted as a renal pelvis calculus at US. **(a)** Longitudinal US image of the left kidney demonstrates a shadowing echogenic focus (cursors) in the renal pelvis that suggests a calculus. **(b)** Transverse nonenhanced CT image reveals atherosclerotic calcification (arrows) of the segmental renal arteries.

TABLE 5
Detection of Any Renal Calculi: Initial Interpretation of 123 Reports

CT Findings	Initial US Interpretation		
	Positive	Negative	Indeterminate
Positive	15	19	5
Negative	2	80	2

Note.—The sensitivity of US was 44% (95% CI: 35%, 53%), and the specificity was 98% (95% CI: 95%, 100%).

The poor sensitivity of US demonstrated in the current study is related to multiple factors, the most important being the excellent contrast resolution of CT that allows discrimination of slight differences in attenuation within the renal pelvis and parenchyma. Helical CT enables acquisition of a volume of data that includes the entire kidney, thus allowing complete evaluation, whereas some portions of the kidney may not be visualized at US. Furthermore, CT is less dependent on factors such as patient body habitus and operator skill that are critical to US. Calculi may be missed at US because of a lack of acoustic shadowing that can occur with intervening tissue of different acoustic impedance. Inappropriate selection of transducer power and focal length can also impair acoustic shadowing (11). Because US has been shown to be sensitive to nonopaque renal calculi, it is unlikely that chemical composition plays a major role in the ability of US to depict calculi (12).

The specificity of US was found to be 90%, equal to that found in the study by Middleton et al (2). In most cases, the source of false-positive findings was unclear after reviewing the corresponding CT scan. However, in two cases, renal arterial calcifications were misinterpreted as renal pelvis calculi. In another case, a ureteral stent easily identified on CT scans was interpreted as a 15-mm renal pelvis calculus on US images. Nonetheless, the relatively high specificity and positive predictive value of US suggests that calculi identified on US images reliably correspond to calculi, as demonstrated with CT, particularly if radiographs are reviewed to exclude vascular calcifications and instrumentation in the collecting system.

As shown by other investigators (2,11, 13), US sensitivity is dependent on calculi size, and our data indicate that US is poor at depicting calculi of 3.0 mm or less. For calculi of this size, the sensitivity of US was found to be 13%. The mean size of missed calculi was 3.3 mm ± 0.6, whereas the mean size of calculi detected with US was 7.1 mm ± 1.2. King et al (11) have shown that the presence of an acoustic shadow depends on the size of a calculus, and it follows that smaller calculi are more likely to be missed if the diagnostic criteria for calculi include acoustic shadowing.

Because approximately 80% of calculi smaller than 5 mm will pass spontaneously (14), it is reassuring that the bulk of missed calculi are relatively small. However, this finding suggests that US is of limited value for evaluating the progression of renal stone disease, in which the identification of new small calculi would be important. Likewise, small fragments that occur after ESWL could be missed.

Although a majority of calculi missed at US were smaller than 4.0 mm, some were much larger, and the sensitivity of US for calculi larger than 7.0 mm was found to be 71%. Indeed, two 15-mm calculi were not visualized at US. The reason for such large calculi not being identified is not clear but may relate to filling of the acoustic shadow from reverberation, a phenomenon previously described by King et al (11).

Accurate determination of calculi size has implications for initial therapy, as well as for follow-up evaluation of nephrolithiasis. In this regard, US is fairly reliable, and the mean difference between CT and US for determining calculus size is minimal (1.5 mm ± 0.7). In a majority of cases (79%), calculi were grouped concordantly on the basis of CT and US measurements. However, in a single case, calculus size was incorrectly categorized as larger than 7.0 mm by using US measurements, potentially resulting in different therapy for this patient. This type of error has similar implications regarding evaluation of calculus fragmentation after ESWL.

Our data indicate that US is a poor modality for demonstrating the full extent of calculi burden. US enabled identification of 39% of the patients with multiple calculi demonstrated at CT, and in only three of these patients were all renal calculi depicted. A study by Vrtiska et al (13) has shown similar difficulties in identifying the full extent of renal calculi with US. Calculi burden and the formation of new calculi is important in the clinical evaluation of patients with renal stone disease, and these findings raise questions concerning the efficacy of renal US for follow-up examination of these patients.

Because improved diagnostic accuracy can be expected if US is observed in real time (15), it is surprising that no substantial difference was found between US sensitivity in the original US reports and retrospective interpretation. This finding probably reflects the fact that our staff members generally perform real-time scanning only in cases in which an abnormality has been initially identified by the sonographer. The improved specificity and positive predictive value of the original interpretation likely results from the opportunity to obtain additional views at the time of scanning, as well as the availability of patient information such as the presence of a ureteral stent, thereby reducing the number of false-positive findings.

The primary limitation of this study was the delay between performance of US and CT. Although the mean time between examinations was less than 1 week, it clearly would be optimal to perform the examinations concurrently to minimize the likelihood of calculi passage prior to the second examination. However, unlike ureteral calculi that are typically being passed when imaging is performed, we think that a majority of renal calculi would not be passed within the interval between examinations used in this study.

Nonenhanced CT should be considered the standard for determining the size, number, and position of renal calculi. At our institution, we have adopted CT as the primary modality for the detection of renal calculi. Although the cost of CT remains a barrier to widespread use, authors of one study (16) suggest that modified nonenhanced CT may in fact be less costly than a combination of radiography and US. The investigators in that study indicate that the radiation dose with a modified helical technique can be reduced to 10 mSv, or approximately that received from three standard radiographs of the abdomen.

References

1. Smith RC, Verga M, McCarthy S, Rosenfield AT. Diagnosis of acute flank pain: value of unenhanced helical CT. AJR Am J Roentgenol 1996; 166:97–101.
2. Middleton WD, Wylie JD, Lawson TL, Foley WD. Renal calculi: sensitivity for detection with US. Radiology 1988; 167:239–244.
3. Levine JA, Neitlich J, Verga M, Dalrymple N, Smith RC. Ureteral calculi in patients with flank pain: correlation of plain radiography with unenhanced helical CT. Radiology 1997; 204:27–31.
4. Yilmaz S, Sindel T, Arslan G, et al. Renal colic: comparison of spiral CT, US and IVU in the detection of ureteral calculi. Eur Radiol 1998; 8:212–217.
5. Boulay I, Holtz P, Foley WD, White B, Begun FP. Ureteral calculi: diagnostic efficacy of helical CT and implications for treatment of patients. AJR Am J Roentgenol 1999; 172:1485–1590.
6. Dawson-Saunders B, Trapp RG. Basic and clinical biostatistics. Norwalk, Conn: Appleton & Lange, 1994; 143–145.
7. Smith RC, Rosenfield AT, Choe KA, et al. Acute flank pain: comparison of non-contrast-enhanced CT and intravenous urography. Radiology 1995; 194:789–794.
8. Amis ES Jr. Epitaph for the urogram. Radiology 1999; 213:639–640.
9. Kaude JV, Williams JL, Wright PG, Bush D, Derau C, Newman RC. Sonographic evaluation of the kidney following extracorporeal shock wave lithotripsy. J Ultrasound Med 1987; 6:299–306.
10. Sommer FG, Jeffrey RB Jr, Rubin GD, et al. Detection of ureteral calculi in patients with suspected renal colic: value of reformatted noncontrast helical CT. AJR Am J Roentgenol 1995; 165:509–513.
11. King W, Kimme-Smith C, Winter J. Renal stone shadowing: an investigation of contributing factors. Radiology 1985; 154:191–196.
12. Pollack HM, Arger PH, Goldberg BB, Mulholland SG. Ultrasonic detection of nonopaque renal calculi. Radiology 1978; 127:233–237.
13. Vrtiska TJ, Hattery RR, King BF, et al. Role of ultrasound in medical management of patients with renal stone disease. Urol Radiol 1992; 14:131–138.
14. Motola JA, Smith AD. Therapeutic options for the management of upper tract calculi. Urol Clin North Am 1990; 17:191–206.
15. Chan V, Hanbidge A, Wilson S, Pron G, Moore L. Case for active physician involvement in US practice. Radiology 1996; 199:555–560.
16. Remer EM, Herts BR, Streem SB, et al. Spiral noncontrast CT versus combined plain radiography and renal US after extracorporeal shock wave lithotripsy: cost-identification analysis. Radiology 1997; 204:33–37.