

Ultrasound is Superior to Computed Tomography for Assessment of Medullary Nephrocalcinosis in Hypoparathyroidism

Alison M. Boyce, Thomas H. Shawker, Suvimol C. Hill, Peter L. Choyke, Michael C. Hill, Robert James, Nancy A. Yovetich, Michael T. Collins, and Rachel I. Gafni

Bone Health Program (A.M.B.), Division of Orthopaedics and Sports Medicine, Children's National Medical Center, Washington, D.C. 20010-2970; Division of Endocrinology and Diabetes (A.M.B.), Children's National Medical Center, Washington, D.C. 20010-2970; Skeletal Clinical Studies Unit (A.M.B., M.T.C., R.I.G.), Clinical Skeletal Diseases Branch, National Institute of Dental and Craniofacial Research, Radiology and Imaging Sciences Branch (T.H.S., S.C.H.), National Institutes of Health, Program in Developmental Endocrinology and Genetics (A.M.B.), Eunice Kennedy Shriver National Institute of Child Health and Human Development, National Institutes of Health, Molecular Imaging Program (P.L.C.), Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20892; Department of Radiology (M.C.H.), George Washington University Hospital, Washington, D.C. 20037; and Rho, Inc (R.J., N.A.Y.), Chapel Hill, North Carolina 27517

Context: Nephrocalcinosis is a complication of hypoparathyroidism and other metabolic disorders. Imaging modalities include ultrasonography (US) and computed tomography (CT). Few studies have compared these modalities, and standard clinical practice is not defined.

Objective: The objective of the study was to determine the preferred method for assessing nephrocalcinosis.

Design: The design of the study was a retrospective, blinded analysis.

Setting: The study was conducted at a clinical research center.

Patients: Twenty-two hypoparathyroid subjects and 7 controls participated in the study.

Interventions: Contemporaneous renal US and CT images were reviewed in triplicate by 4 blinded radiologists. Nephrocalcinosis was classified using a 0–3 scale with 0 meaning no nephrocalcinosis and 3 meaning severe nephrocalcinosis.

Main Outcome Measures: Intraobserver, interobserver, and interdevice agreements were measured.

Results: Intraobserver agreement was high, with an overall weighted kappa of 0.83 for CT and 0.89 for US. Interobserver agreement was similar between modalities, with kappas of 0.74 for US and 0.70 for CT. Only moderate agreement was found between US and CT scores, with an intermodality kappa of 0.47 and 60% concordance. Of discordant pairs, 81% had higher US scores and only 19% had higher CT scores. Of nephrocalcinosis seen on US and not CT, 45%, 46%, and 9% were grades 1, 2, and 3, respectively. Overall, US scores were higher than CT with a cumulative odds ratio (95% confidence interval) of 5.97 (2.60, 13.75) ($P < .01$). In controls, 100% of US ratings were 0, and 95% of CT ratings were 0.

Conclusions: US is superior to CT for assessment of mild to moderate nephrocalcinosis in patients with hypoparathyroidism. This finding, in combination with its low cost, lack of radiation, and portability, defines US as the preferred modality for assessment of nephrocalcinosis. (*J Clin Endocrinol Metab* 98: 989–994, 2013)

Renal calcifying disorders include nephrocalcinosis, describing parenchymal calcium deposition (1–3), and nephrolithiasis, defined as calcifications within the collecting system (4). Despite anatomically different locations, these conditions share common etiological features. Unlike nephrolithiasis, which may occur in healthy individuals, nephrocalcinosis suggests an underlying endocrine or metabolic disorder (2, 5), such as hypoparathyroidism. Deficiency or decreased activity of parathyroid hormone results in hypocalcemia, hyperphosphatemia, and relative hypercalciuria, placing patients at risk for renal calcification (6–9). Unlike nephrolithiasis, nephrocalcinosis is asymptomatic, and severe disease may progress to renal insufficiency (6, 9). Early diagnosis is important because therapeutic interventions may stabilize, slow or, rarely, reverse disease (10, 11). Thus, surveillance in high-risk patients is a crucial component of management.

Diagnostic imaging for nephrocalcinosis includes radiographs, computed tomography (CT), and ultrasonography (US). Radiographs are insensitive due to poor delineation of renal anatomy and confounding by bowel gas. CT offers superior contrast resolution and definition of anatomic structures and is the gold standard for nephro-

lithiasis (12). US, however, has advantages for assessment of nephrocalcinosis because it is radiation-free, portable, and relatively inexpensive (5). Few studies have systematically compared these modalities, and best clinical practice is undefined.

This study compares renal US and CT in subjects with hypoparathyroidism to determine the preferred imaging modality for assessment of nephrocalcinosis.

Materials and Methods

Subjects

Contemporaneous renal US and CT images were obtained from 22 subjects with hypoparathyroidism. Subjects ranged in age from 9 to 50 years, with a median age of 32.5 years. Table 1 includes demographic and laboratory data for each subject at the time imaging was obtained. All subjects were initiated on calcium and vitamin D analogs at the time of diagnosis, with multiple dose adjustments over the ensuing years by their health care providers. At the time imaging was obtained at the National Institutes of Health (NIH), 7 of the subjects were being managed with sc injections of PTH 1–34. Subjects were thus treated with a range of therapeutic regimens, which varied over their disease course.

Table 1. Subject Demographics and Clinical Data

Subject	Sex	Age, y	Diagnosis	Duration of Disease, y	Serum Ca, mg/dL	Urine Ca, mg/kg · d	eGFR, mL/min per 1.73 m ²	Mean US Grade	Mean CT Grade	NL
1	F	9	CaSR mutation	9	7.8	7.45	214	2	0	N
2	F	9	CaSR mutation	9	10.0	6.49	79	3	2.67	N
3	F	9	CaSR mutation	9	9.4	4.14	81	2.17	2.83	N
4	M	12	CaSR mutation	12	8.4	1.79	120	3	2.67	N
5	M	15	AIRE mutation	13	8.8	4.65	117	2.67	1.33	N
6	F	16	CaSR mutation	16	8.0	3.61	49	2.33	0.83	Y
7	M	16	CaSR mutation	16	8.4	6.56	113	0.33	0	N
8	F	18	22q11 deletion	18	8.1	5.81	108	0.50	0.33	N
9	M	18	AIRE mutation	9	8.9	3.64	115	0	0	N
10	F	29	Postsurgical	5	7.8	3.89	92	0.83	0	N
11	M	31	Congenital idiopathic	31	7.1	2.72	70	1.17	0	N
12	F	34	Postsurgical	5	8.5	7.88	59	0	0	N
13	F	38	Postsurgical	6	8.1	3.59	114	2	0	N
14	M	40	CaSR mutation	40	8.3	10.89	72	3	3	Y
15	F	40	22q11 deletion	40	7.6	3.74	97	0.33	0.17	N
16	F	41	Postsurgical	6	7.8	1.14	103	0	0.83	N
17	F	42	Postsurgical	4	7.5	3.18	87	0.33	0	N
18	F	42	Postsurgical	4	9.2	3.60	101	0	0	N
19	F	44	Postsurgical	8	7.8	3.41	63	2.5	0.83	N
20	F	45	Postsurgical	4	7.6	5.96	105	0.33	0	N
21	M	49	Acquired idiopathic	5	7.3	2.42	81	0.5	0	N
22	M	50	CaSR mutation	50	7.4	8.90	70	0	0	N

Abbreviations: AIRE, autoimmune regulator; Ca, calcium; CaSR, calcium sensing receptor; eGFR, estimated glomerular filtration rate; F, female; M, male; N, no; NL, nephrolithiasis; Y, yes. Normal range for serum calcium is 8.2–10 mg/dL; to convert serum calcium to SI units (millimoles per liter), divide by 4. Normal urine calcium excretion is less than 4 milligrams per kilogram per day; to convert SI units (millimoles per kilograms per day), divide by 40. The estimated glomerular filtration rate was determined using the Schwartz equation for patients under age 18 years (22), and the chronic kidney disease epidemiology collaboration (CKD-EPI) equation for subjects aged 18 years and older (23). Classification of the stages of chronic kidney disease based on the Kidney Disease Outcomes Quality Initiative guidelines are as follows: eGFR 90 or greater is stage 1; eGFR 60–89 is stage 2; eGFR 30–59 is stage 3; eGFR 15–29 is stage 4; eGFR less than 15 or dialysis is stage 5 (24).

Seven controls without known risk factors for nephrocalcinosis (aged 30–74 years, median 45.8 years) were identified using NIH's Biomedical Translational Research Information System (13) with approval of the Office of Human Subjects Research. Subjects participated in institutional review board-approved protocols at the NIH and gave informed consent/assent.

US and CT methodology and scoring system

Renal US was performed using a curved array sector with a 2- to 4-MHz frequency. Subjects underwent multidetector helical noncontrast CT. Slice thickness was variable: 2 mm in 8 subjects, 2.5 mm in 6 subjects, 5 mm in 14 subjects, and 10 mm in 1 control subject.

Each scan was copied in triplicate, yielding a total of 87 studies per modality. Studies were deidentified, assigned unique identification numbers, and randomly sorted. Two radiologists specializing in US (T.H.S. and M.C.H.) and 2 specializing in CT (S.C.H. and P.L.C.) interpreted their respective modalities. Thus, within each subject there were 6 replicate US scores (2 radiologists \times 3 readings) and 6 replicate CT scores, for 174 (6 scores \times 29 subjects) replicate readings per modality. Radiologists were blinded to diagnosis and were not informed that they were reading identical studies in triplicate.

Studies were evaluated for both nephrocalcinosis and nephrolithiasis. Nephrocalcinosis was defined as the presence of calcifications identified within the renal parenchyma, outside the collecting system. Nephrolithiasis was defined as the presence of calcifications located within the collecting system.

A nephrocalcinosis grading system was developed and agreed upon by all radiologists prior to analysis, based on previously described radiographic features of progressive nephrocalcinosis (1, 3). US grades were as follows: grade 0, no echogenicity; grade 1, mild echogenicity around medullary pyramid borders; grade 2, moderate echogenicity around and inside pyramids; and grade 3, severe echogenicity of entire pyramids. For CT the grades were as follows: grade 0, no calcific deposits; grade 1, 1–3 pyramidal punctate calcifications; grade 2, increased pyramidal density; grade 3, calcification of pyramids (Figure 1). The presence of any involved papillae yielded a score greater than 0; no minimal number of papillae was required. When features of multiple

grades were present, or the extent of disease differed between the 2 kidneys, the higher score was assigned.

Statistical analysis

Because scores were on an ordinal rather than binomial scale, the weighted kappa statistic was used to summarize intraobserver, interobserver, and intermodality agreements. The weighted kappa statistic adjusts for disagreement on the ordinal scale (eg, score of 3 vs 1 or 2 vs 1) by using weights proportional to the degree of difference in ordinal values. A kappa coefficient of 1 implies perfect agreement; the level of agreement decreases with a decreasing score. Kappa statistics were complicated by replicate scores; thus, all possible pair-wise comparisons within and between subject-level replicates of different raters were included in calculation of weighted kappa statistics. This approach was permissible due to perfect data balance among subjects. Because standard formulas for standard errors of kappa are not valid for replicated measures, bootstrap methodology with 95th percentile confidence limits was used to estimate the error in weighted kappa.

Spearman rank correlations were performed to test for any monotonic increasing or decreasing associations between duration of hypoparathyroidism and the mean nephrocalcinosis score (across both CT and US for each subject). Next, the duration of hypoparathyroidism was dichotomized among subjects as 1, meaning short duration of less than 10 years, and 2, meaning long duration with 10 or more years of hypoparathyroidism (the split was based on a 9 year median duration). The weighted kappa statistic testing for agreement between CT and US was then estimated separately for subjects with the short or long duration of hypoparathyroidism.

Generalized linear mixed models (GLMM) (SAS, version 9.2, Proc GLIMMIX; SAS Institute, Cary, North Carolina) were fit to the data using a multinomial distribution using with a cumlogit link (proportional odds), Laplace parameter estimation, and random effects for subject and specialist. To improve the model convergence, the GLMM models were fit to a 3-point nephrocalcinosis scale (grades 1 and 2 combined). CT thickness was dichotomized as thin and thick CT slices. The GLMM was fit with and without CT thickness as a model covariate.

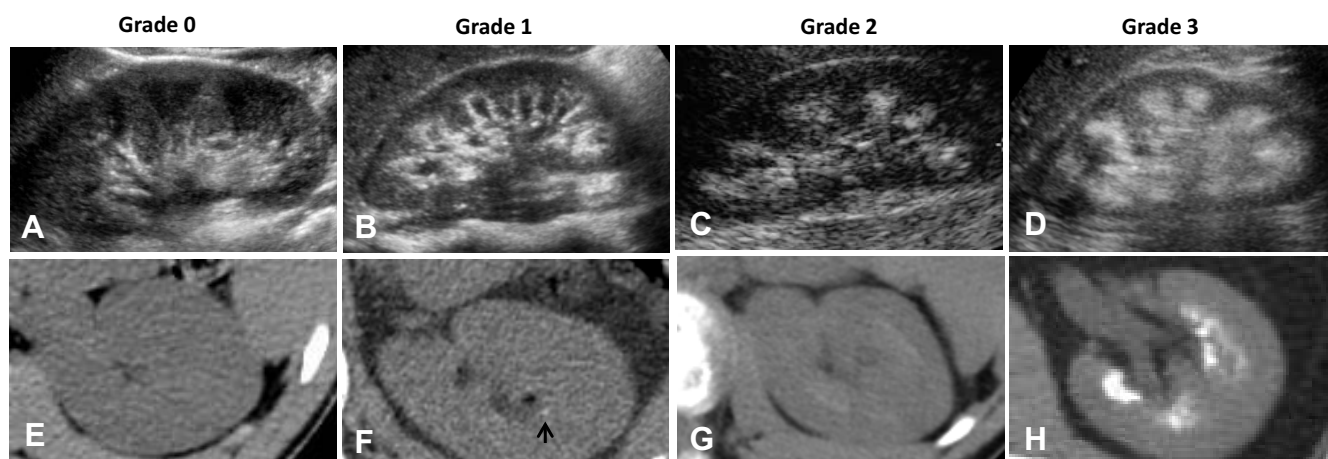


Figure 1. Representative images for nephrocalcinosis grading system. Panels A–D are US images corresponding to grades 0, 1, 2, and 3, respectively. Panels E–H are CT images corresponding to grades 0, 1, 2, and 3 respectively. The arrow in panel F points to a single punctate calcification in the medullary periods typical of grade 1 nephrocalcinosis.

Results

Intra- and interobserver agreement was high for both modalities. Pair-wise comparisons between raters were in complete agreement for 81.6% of the CT scores and 69.7% of the US scores.

Only moderate agreement was found between US and CT scores, with a weighted kappa of 0.47 (0.221, 0.663), and 60% concordance (Figure 2). Of the 40% discordant pairs, 81% had higher US scores and only 19% had higher CT scores (Figure 2). Eighty of 174 replicate US scores (46%) were greater than 0, whereas only 47 of 174 replicate CT scores (27%) were greater than 0. Of nephrocalcinosis seen on US and not CT, 45%, 46%, and 9% were grades 1, 2, and 3, respectively. Overall, US scores were higher than CT with a cumulative odds ratio (OR) of 5.97 (2.60, 13.75) ($P < .01$). In 7 control subjects, 100% of US ratings were 0, whereas 95% of CT ratings were 0.

The weighted kappa improved to 0.68 (0.375, 0.861) when agreement with US was limited to just the thick CT slices (5 and 10 mm). However, agreement was found to be very poor when the comparison was limited to the US and thin CT slices (2 and 2.5 mm), with a weighted kappa of only 0.08 (−0.151, 0.246). Adding covariates for CT thickness or duration of hypoparathyroidism had no significant effect on the cumulative OR (effect of CT thickness: $P = .5607$; effect of duration: $P = .7156$).

A follow-up analysis was subsequently performed in 1 hypoparathyroid subject. The initial US and CT revealed

no nephrocalcinosis; however, after a 12-month period, he developed grade 1 nephrocalcinosis on US, whereas CT showed no nephrocalcinosis.

Discussion

Analysis of contemporaneous renal US and CT revealed that US is superior to CT for the detection of nephrocalcinosis in hypoparathyroidism, particularly in the early stages of the disease. Nephrocalcinosis screening is an important component of management in patients with hypoparathyroidism, who are at risk due to impairment in urinary calcium reabsorption (6, 8, 14). Development of nephrocalcinosis is somewhat unpredictable, but is likely related to both the degree and duration of hypercalciuria. Urinary calcium excretion may vary widely between and within individuals and is highly dependent on treatment, medication compliance, and diet. Etiology of hypoparathyroidism may also play a role in the development of nephrocalcinosis, as hypercalciuria is typically more pronounced in patients with activating mutations of the calcium sensing receptor (15). In addition to hypoparathyroidism, nephrocalcinosis is seen in a variety of endocrine and metabolic disorders, most frequently associated with hypercalciuria, hyperoxaluria, hyperphosphaturia (5), and prematurity (16). Thus, determining the most accurate modality to identify nephrocalcinosis in its earliest stages is important to a broad patient population.

Defining evidence-based practice for the evaluation of nephrocalcinosis has been limited by a lack of studies comparing imaging modalities. In 1990, Manz et al (17) evaluated 12 patients with nephrocalcinosis (including 9 biopsy proven cases) using radiographs, US, and CT. CT and x-ray showed concordance with histopathology, but US findings were nonspecific. Kim et al (18) performed radiographs, US, and CT in 18 long-term furosemide users and found nephrocalcinosis on 15 ultrasounds, 12 CT scans, and 2 radiographs. Cramer et al (19) performed a prospective comparison of US and CT in rabbits, in which nephrocalcinosis was induced and radiographic findings were correlated with histopathology. US demonstrated superior sensitivity (96% vs 64% for CT) but lower specificity for nephrocalcinosis (85% vs 96%). More recently, Cheidde et al (20) compared US, CT, and x-ray in 62 patients with nephrocalcinosis; however, interpretation was limited by low inter- and intraobserver agreement.

Similar to the studies in rabbits and furosemide users, our results suggest US is more sensitive than CT for the evaluation of nephrocalcinosis. US appeared particularly sensitive for mild to moderate nephrocalcinosis, as most discordant scores were US grades 1 and 2 paired with

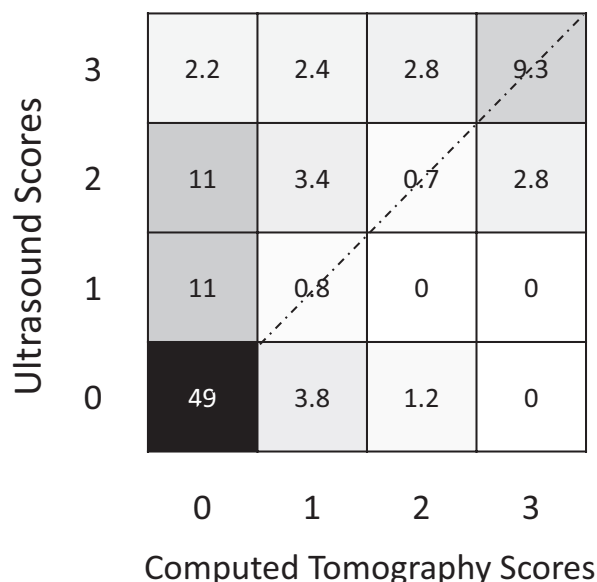


Figure 2. Density plot showing relative frequencies of paired observations of US and CT scores, expressed as percentages. The cells along the diagonal show the relative frequency of the scores in which both devices are in perfect agreement. The cells above the diagonal have a higher density than the cells below the diagonal, illustrating the tendency of the US to score individual scans higher than CT.

negative CT studies. These findings suggest that early nephrocalcinosis becomes visible on US prior to CT. This is supported by observations in a subject with initially normal studies, who went on to develop grade 1 nephrocalcinosis on US that was not seen on CT. Longitudinal studies are needed to investigate the radiological evolution of nephrocalcinosis.

The physics underlying the inferiority of CT to detect early of nephrocalcinosis is not entirely understood. Detection of calcification on CT requires sufficient density to attenuate x-rays transiting the body. The density of calcification early in the course of nephrocalcinosis may not be sufficient for detection by CT. In contrast, US depends on the ability of small reflectors, such as small calcific aggregations with significant impedance mismatches to normal tissue, to act as strong scatterers of ultrasound waves, resulting in enhanced detection of the small calcium deposits found in nephrocalcinosis (21).

There are several inherent study limitations. CT slice thicknesses were not uniform between subjects. There was greater agreement between US scores and thick-slice CT as compared with thin-slice CT scores; however, the odds of scoring lower on CT was not affected by slice thickness. Conventional ultrasound transducers (2–4 MHz) optimized for abdominal imaging were used in this study. It is possible that higher frequency transducers (7–9 MHz) may have resulted in a higher detection rate of nephrocalcinosis; however, it is not always possible to visualize the kidneys with these higher frequency transducers. Higher-frequency transducers resulted in lower specificity for US compared with CT in 1 study (19). Finally, a central limitation is the lack of a gold standard for diagnosis because renal biopsies were not available to confirm the presence or absence of nephrocalcinosis. For the purposes of this study, the assumption was made that positive findings were true positives in this high-risk population. However, if US is less specific for nephrocalcinosis, as was reported in the animal study by Cramer et al (19), cases seen on US and not CT may represent false positives. That no nephrocalcinosis was observed in controls argues against this possibility. Even if this were the case, for clinical management purposes a more sensitive modality may be preferable to one that is less sensitive but more specific, given that: 1) nephrocalcinosis may stabilize or even reverse with improved clinical management, and therefore, patients who are undiagnosed lose the opportunity to prevent disease progression; 2) treatment is conservative and involves optimizing therapy to avoid hypercalciuria, and thus, a false-positive finding is unlikely to result in invasive or costly interventions; and 3) US offers practical and safety benefits over CT, including lack of ionizing radiation, lower cost, and portability. The possibility of overdiagnosis by a

potentially less sensitive imaging modality must thus be weighed against the additional cost and risk associated with CT.

In conclusion, ultrasound is superior to CT for the detection of mild to moderate nephrocalcinosis in hypoparathyroidism. This finding is likely generalizable to other diseases associated with nephrocalcinosis. Given its suggested greater sensitivity, lack of radiation, portability, and lower cost, US is the preferred modality for evaluation of nephrocalcinosis.

Acknowledgments

We acknowledge Shawn Thomas and James Cimino for their assistance with this project.

Address all correspondence and requests for reprints to: Alison M. Boyce, MD, Division of Endocrinology and Diabetes, Children's National Medical Center, 111 Michigan Avenue NW, Washington, D.C. 20010-2970. E-mail: aboyce@childrensnational.org

This work was supported by the Division of Intramural Research, National Institute of Dental and Craniofacial Research, Eunice Kennedy Shriver National Institute of Child Health and Human Development, National Institutes of Health, Department of Health and Human Services, and the Bone Health Program, Children's National Medical Center.

Disclosure Summary: All authors state that they have no conflicts of interest.

References

1. al-Murrani B, Cosgrove DO, Svensson WE, Blaszczyk M. Echogenic rings—an ultrasound sign of early nephrocalcinosis. *Clin Radiol*. 1991;44:49–51.
2. Sayer JA, Carr G, Simmons NL. Nephrocalcinosis: molecular insights into calcium precipitation within the kidney. *Clin Sci (London, England)*. 2004;106:549–561.
3. Patriquin H, Robitaille P. Renal calcium deposition in children: sonographic demonstration of the Anderson-Carr progression. *AJR Am J Roentgenol*. 1986;146:1253–1256.
4. Evan AP. Physiopathology and etiology of stone formation in the kidney and the urinary tract. *Pediatr Nephrol (Berlin, Germany)*. 2010;25:831–841.
5. Habbig S, Beck BB, Hoppe B. Nephrocalcinosis and urolithiasis in children. *Kidney Int*. 2011;80:1278–1291.
6. Pearce SH, Williamson C, Kifor O, et al. A familial syndrome of hypocalcemia with hypercalciuria due to mutations in the calcium-sensing receptor. *N Engl J Med*. 1996;335:1115–1122.
7. Shoback D. Clinical practice. Hypoparathyroidism. *N Engl J Med*. 2008;359:391–403.
8. Lienhardt A, Bai M, Lagarde JP, et al. Activating mutations of the calcium-sensing receptor: management of hypocalcemia. *J Clin Endocrinol Metab*. 2001;86:5313–5323.
9. Sorheim JI, Husebye ES, Nedrebo BG, et al. Phenotypic variation in a large family with autosomal dominant hypocalcaemia. *Horm Res Paediatr*. 2010;74:399–405.
10. Auron A, Alon US. Resolution of medullary nephrocalcinosis in children with metabolic bone disorders. *Pediatr Nephrol (Berlin, Germany)*. 2005;20:1143–1145.

11. Curry NS, Gordon L, Gobien RP, Lott M. Renal medullary “rings”: possible CT manifestation of hypercalcemia. *Urol Radiol*. 1984;6: 48–50.
12. Mandeville JA, Gnessin E, Lingeman JE. Imaging evaluation in the patient with renal stone disease. *Semin Nephrol*. 2011;31:254–258.
13. Cimino JJ, Ayres EJ. The clinical research data repository of the US National Institutes of Health. *Stud Health Technol Inform*. 2010; 160:1299–1303.
14. Raue F, Pichl J, Dorr HG, et al. Activating mutations in the calcium-sensing receptor: genetic and clinical spectrum in 25 patients with autosomal dominant hypocalcaemia—a German survey. *Clin Endocrinol (Oxf)*. 2011;75:760–765.
15. Yamamoto M, Akatsu T, Nagase T, Ogata E. Comparison of hypocalcemic hypercalciuria between patients with idiopathic hypoparathyroidism and those with gain-of-function mutations in the calcium-sensing receptor: is it possible to differentiate the two disorders? *J Clin Endocrinol Metab*. 2000;85:4583–4591.
16. Schell-Feith EA, Kist-van Holthe JE, van der Heijden AJ. Nephrocalcinosis in preterm neonates. *Pediatr Nephrol (Berlin, Germany)*. 2010;25:221–230.
17. Manz F, Jäschke W, van Kaick G, Waldherr R, Willich E. Nephrocalcinosis in radiographs, computed tomography, sonography and histology. *Pediatr Radiol*. 1980;9:19–26.
18. Kim YG, Kim B, Kim MK, et al. Medullary nephrocalcinosis associated with long-term furosemide abuse in adults. *Nephrology, Dialysis, Transplantation*. 2001;16:2303–2309.
19. Cramer B, Husa L, Pushpanathan C. Nephrocalcinosis in rabbits—correlation of ultrasound, computed tomography, pathology and renal function. *Pediatr Radiol*. 1998;28:9–13.
20. Cheidde L, Ajzen SA, Tamer Langen CH, Christophalo D, Heilberg IP. A critical appraisal of the radiological evaluation of nephrocalcinosis. *Nephron Clin Pract*. 2007;106:c119–c124.
21. Zagzebski JA. *Essentials of Ultrasound Physics*; 1996:72–88.
22. Schwartz GJ, Munoz A, Schneider MF, et al. New equations to estimate GFR in children with CKD. *J Am Soc Nephrol*. 2009;20: 629–637.
23. Levey AS, Stevens LA, Schmid CH, et al. A new equation to estimate glomerular filtration rate. *Ann Intern Med*. 2009;150:604–612.
24. K/DOQI clinical practice guidelines for chronic kidney disease: evaluation, classification, and stratification. *Am J Kidney Dis*. 2002; 39(2 suppl 1):S1–S266.



Make sure your patients are getting the best medical care. Learn more about The Endocrine Society's ***The Evaluation of Thyroid Nodules Practice Improvement Module(PIM)***.

www.endoselfassessment.org