UCLA UCLA Previously Published Works

Title

Prospective Validation of Two 4D-CT-Based Scoring Systems for Prediction of Multigland Disease in Primary Hyperparathyroidism

Permalink https://escholarship.org/uc/item/68v8t3ct

Journal American Journal of Neuroradiology, 37(12)

ISSN 0195-6108

Authors

Sho, S Yilma, M Yeh, MW <u>et al.</u>

Publication Date 2016-12-01

DOI

10.3174/ajnr.a4948

Peer reviewed

Prospective Validation of Two 4D-CT–Based Scoring Systems for Prediction of Multigland Disease in Primary Hyperparathyroidism

¹⁰S. Sho, ¹⁰M. Yilma, ¹⁰M.W. Yeh, ¹⁰M. Livhits, ¹⁰J.X. Wu, ¹⁰J.K. Hoang, and ¹⁰A.R. Sepahdari

ABSTRACT

BACKGROUND AND PURPOSE: Patients with multigland primary hyperparathyroidism are at higher risk for missed lesions on imaging and failed parathyroidectomy. The purpose of this study was to prospectively validate the ability of previously derived predictive score systems, the composite multigland disease score, and the multiphase multidetector contrast-enhanced CT (4D-CT) composite multigland disease score, to identify patients with a high likelihood of multigland disease.

MATERIALS AND METHODS: This was a prospective study of 71 patients with primary hyperparathyroidism who underwent 4D-CT and successful parathyroidectomy. The size and number of lesions identified on 4D-CT, serum calcium levels, and parathyroid hormone levels were collected. A composite multigland disease score was calculated from 4D-CT imaging findings and the Wisconsin Index (the product of the serum calcium and parathyroid hormone levels). A 4D-CT multigland disease score was obtained by using the CT data alone.

RESULTS: Twenty-eight patients with multigland disease were compared with 43 patients with single-gland disease. Patients with multigland disease had a significantly smaller lesion size (P < .01) and a higher likelihood of having either ≥ 2 or 0 lesions identified on 4D-CT (P < .01). Composite multigland disease scores of ≥ 4 , ≥ 5 , and 6 had specificities of 72%, 86%, and 100% for multigland disease, respectively. 4D-CT multigland disease scores of ≥ 3 and 4 had specificities of 74% and 88%.

CONCLUSIONS: Predictive scoring systems based on 4D-CT data, with or without laboratory data, were able to identify a subgroup of patients with a high likelihood of multigland disease in a prospectively accrued population of patients with primary hyperparathyroidism. These scoring systems can aid in surgical planning.

 $\label{eq:ABBREVIATIONS: 4D-CT = multiphase multidetector contrast-enhanced CT; MGD = multigland disease; PTH = parathyroid hormone; SGD = single-gland disease; WIN = Wisconsin Index$

Primary hyperparathyroidism is characterized by excessive parathyroid hormone production resulting in hypercalcemia. Surgical resection remains the only definitive cure. In recent years, preoperative localization of the abnormal parathyroid gland is routinely performed and plays an integral part in operative guidance.¹⁻³ Most cases of primary hyperparathyroidism are caused by a single parathyroid adenoma; however, 10%–30% of patients are known to have multigland disease (MGD).^{4.5} These patients pose a considerable challenge in preoperative localization

http://dx.doi.org/10.3174/ajnr.A4948

because the performance of technetium Tc99m sestamibi scanning and sonography is inferior in patients with MGD compared with patients with single-gland disease (SGD).^{6,7} The decision to perform minimally invasive parathyroidectomy or 4-gland exploration relies heavily on preoperative imaging findings. Accurate preoperative identification of abnormal parathyroid lesions promotes effective operative planning and patient counseling, and can prevent a failed operation.^{3,8-10} For example, reliable preoperative determination of SGD allows a surgeon to perform minimally invasive unilateral parathyroidectomy, which minimizes the incision length and reduces the risk of bilateral recurrent laryngeal nerve injury.

Multiphase multidetector contrast-enhanced CT (4D-CT) has emerged in recent years as a new technique of preoperative localization of an abnormal parathyroid gland. 4D-CT has been shown to successfully localize abnormal parathyroid glands that have been missed by scintigraphy and sonography. Furthermore, multiple studies have shown the superior sensitivity of 4D-CT in identifying abnormal parathyroid glands in patients with both SGD

Received April 3, 2016; accepted after revision June 11.

From the Department of Surgery (S.S., M.Y., M.W.Y., M.L., J.X.W.), Section of Endocrine Surgery, and Department of Radiological Sciences (A.R.S.), David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, California; and Department of Radiology (J.K.H.), Duke University Medical Center, Durham, North Carolina.

Please address correspondence to Ali R. Sepahdari, MD, Department of Radiological Sciences, David Geffen School of Medicine, University of California, Los Angeles, 757 Westwood Plaza, Suite 1621D, Los Angeles, CA 90095; e-mail: ali.sepahdari@ gmail.com; @alisepahdari

and MGD.⁶⁻⁸ However, the sensitivity of 4D-CT for MGD remains low at 32%–53%, compared with its sensitivity for SGD (88%-93%).⁶⁻⁸

To improve identification of patients with a high likelihood of MGD, Sepahdari et al¹¹ developed a composite MGD score based on 4D-CT imaging and biochemical data derived from a retrospective review of 155 patients. Variables included in this scoring system were the following: 1) the size of the largest lesion identified on 4D-CT, 2) the number of lesions identified on 4D-CT, and 3) the Wisconsin Index (product of serum calcium and parathyroid hormone levels). Scores of ≥ 4 , ≥ 5 , and 6 had specificities of 81%, 93%, and 98%, respectively, for the identification of MGD.¹¹ Although promising, this MGD scoring system was derived solely from a retrospective review and therefore was limited due to biases inherent to its study design. The aim of the current study was to evaluate the performance of the MGD scoring system for predicting MGD by applying it to a prospectively accrued patient population.

MATERIALS AND METHODS

Study Subjects

Institutional review board approval was obtained for this study, which was performed with a waiver of informed consent and a waiver of Health Insurance Portability and Accountability Act authorization. All subjects having 4D-CT for primary hyperparathyroidism were accrued in a prospective manner during the 12 months between January 2014 and January 2015 in a single academic institution (University of California Los Angeles). Subjects were identified at the time of the initial scan. Clinical data and CT-derived data were recorded. Periodic chart review was then performed to track clinical outcomes. Subjects who went on to have successful parathyroidectomy, defined as an intraoperative parathyroid hormone (PTH) drop of 50% and/or at least 6 months of postoperative eucalcemia, were included for further analysis.

4D-CT Technique

Imaging was performed on either a 64–detector row scanner (Somatom Definition; Siemens, Erlangen, Germany) or a 256–detector row scanner (Somatom Definition Flash; Siemens). Scanning included noncontrast, arterial phase, and delayed-phase images from the hard palate to the carina. The parameters for all 3 phases were the following: section thickness, 0.6 mm; tube rotation time, 0.5 seconds; pitch factor, 1; FOV, 24 cm; 120 kVp; 230 reference mAs with automated tube current modulation (CARE Dose4D; Siemens). Arterial phase images were obtained 25 seconds following the initiation of a 100- to 120-mL IV bolus of iohexol, 350 mg of iodine/mL, injected through either a 20- or 22-ga antecubital catheter at either 4 or 3 mL/s. The delayed phase was acquired 30 seconds after the arterial phase ended. All images were reconstructed at 1-mm section thickness in the axial, coronal, and sagittal planes and reviewed in the PACS.

Lesion Localization

All parathyroid lesions were classified as correctly or incorrectly localized on 4D-CT by correlating the operative notes with the original radiology reports and by using anatomic landmarks re-

Table 1: MGD scores^a

gle lesion: 0 Itiple lesions: 2 lesions: 2
lesions: 2
s mm: 0 3 mm: 1 mm or no lesions: 2
500: 0

^a The composite MGD score is based on all 3 components in Table 1 and ranges from 0 to 6. The 4D-CT MGD score does not include the WIN and ranges from 0 to 4.

ported in both the operative and radiology reports. Lesions were described by using a system based on the anteroposterior location relative to the course of the recurrent laryngeal nerve. Lesions posterior to the expected course of the nerve were defined as superior parathyroid glands, and those anterior to the recurrent laryngeal nerve were defined as inferior parathyroid glands. The location along a superior-inferior axis was also described with the thyroid isthmus and the lower edge of the thyroid gland as landmarks. Lesions lying outside these typical locations were described as ectopic. Radiology reports were generated by a subspecialty-certified neuroradiologist with 10 years' experience in CT interpretation, including 5 years' experience in interpretation of 4D-CT. Sensitivities for lesion localization were based on these original radiology reports.

Predictors of Multigland Disease

4D-CT imaging and biochemical predictors of MGD were originally proposed by Sepahdari et al¹¹ on the basis of prior surgical literature.^{12,13} 4D-CT imaging predictors were the number of lesions identified on the original radiology report and the size of the largest lesion (maximum diameter in any plane). Laboratory data included serum calcium levels (milligram/deciliter), serum PTH levels (picogram/milliliter), and the Wisconsin Index (WIN). The WIN is the product of the serum calcium levels (milligram/deciliter) and PTH levels (picogram/milliliter) and was shown to help discriminate MGD and SGD in prior studies.^{11,12}

A composite MGD score was derived from variables of lesion size on 4D-CT, the number of prospectively detected lesions on 4D-CT, and the WIN. Each variable contributed up to 2 points to the MGD scores (Table 1). The cutoff values used to assign points in the score were determined in a previous study by Sepahdari et al,¹¹ on the basis of prior literature for lesion size and ranges of biochemical markers. Maximum lesion size of >13 mm, 7-13 mm, and <7 mm were assigned scores of 0, 1, and 2. A single prospectively identified candidate lesion was assigned a score of zero versus a score of 2 for multiple candidate lesions or no candidate lesion. WINs of >1600, 800-1600, and <800 were assigned scores of 0, 1, and 2, respectively. A second scoring system, the 4D-CT MGD score, was based only on the 4D-CT imaging variables of lesion size and the number of prospectively detected lesions on 4D-CT. The composite MGD score ranged from 0 to 6. The 4D-CT MGD score ranged from 0 to 4. For both scoring systems, a higher score more strongly favored MGD.

Table 2: Characteristics of patients with MGD and SGD^a

	All Subjects	MGD	SGD	P Value
Patient characteristics				
No. of patients	71	28	43	
No. of glands	120	77	43	
Mean age (yr)	62 (range, 21–88)	62.4	62.4	.99
Female	58 (82%)	26 (93%)	32 (74%)	.15
4D-CT characteristics				
Mean size of abnormal glands (mm)	9.6 (6.2)	7.5 (3.3)	11.2 (7.4)	.003
Median size of abnormal glands (mm)	9	7	10	
Preoperatively detected lesions				
1	33 (46%)	6 (21%)	27 (63%)	.003
≥2	28 (39%)	16 (58%)	12 (28%)	
None	10 (14%)	6 (21%)	4 (9.3)	
Biochemical markers				
Serum calcium level (mg/dL)	10.6 (0.6)	10.4 (0.6)	. ,	.08
Serum parathyroid hormone level (pg/mL)	93 (43)	87 (41)	97 (44)	.34
WIN	989 (484)	910 (453)	1040 (501)	.27
MGD scores				
Composite MGD score	3.4 (1.6)	4.3 (1.5)	2.7 (1.4)	<.001
4D-CT MGD score	2.1 (1.4)	2.9 (1.1)	1.5 (1.5)	<.001
4D-CT sensitivity				
Detection of lesions	82 (69%)	43 (56%)		
Detection of all lesions in individual patients	47 (67%)	8 (29%)	39 (91%)	
Parathyroid adenoma localization				
\geq 1 Incorrectly localized or missed adenomas	· · ·	20 (71%)	. ,	
All adenomas localized correctly	47 (67%)	8 (29%)	39 (91%)	

^a Values are expressed as mean (SD) except as specified.

Data Analysis

The characteristics of MGD were compared with those of SGD for individual variables and the scoring system. The χ^2 test was used to assess differences in patients with MGD and SGD for categoric data. The Student *t* test was used to assess differences between MGD and SGD for continuous variables. Receiver operating characteristic analysis was performed to determine the sensitivity and specificity of each feature for predicting MGD. P < .05 was the threshold used for statistical significance for all tests. The data were entered into an Excel spreadsheet (2007 Version; Microsoft, Redmond, Washington). Statistical analyses were performed by using SAS Enterprise Guide (Version 4.2; SAS Institute, Cary, North Carolina).

RESULTS

Study Subjects

One-hundred four patients with primary hyperparathyroidism were imaged with 4D-CT during the study period from January 2014 to January 2015. Twenty-nine patients had no record of parathyroidectomy and therefore were excluded due to an inability to definitively categorize them as having either SGD or MGD on the basis of operative diagnoses. Two had failed parathyroidectomy, 1 had parathyroid carcinoma, and 1 patient died of an unrelated cause before the operation. The final analysis group included 71 patients (Table 2). Failed parathyroidectomies were defined as a lack of persistent eucalcemia for 6 months postoperatively and likely indicate that the abnormal parathyroid adenoma responsible for hyperparathyroidism was not removed in these patients. Therefore, it is unclear whether these patients had a single abnormal adenoma that was missed or had MGD and not all of the abnormal adenomas were removed. Therefore, these 2 patients were removed from analysis because we could not definitively place them in either the SGD or MGD group. Of these

patients, 28 had MGD and 43 had SGD. There were 77 abnormal glands among the 28 patients with MGD, with 120 abnormal glands identified surgically. There were no statistically significant differences between the SGD and MGD groups with regard to age and sex.

Lesion Localization

Among the 28 patients with MGD, 43 of 77 abnormal glands (56%) were identified preoperatively by using 4D-CT. For 8 of 28 patients (29%), all of the abnormal glands in an individual patient could be identified preoperatively with 4D-CT alone. Of the 43 patients with SGD, 39 (91%) lesions were identified preoperatively with 4D-CT.

Imaging Findings and Biochemical Factors

Patients with MGD had a smaller mean lesion size of 7.5 mm compared with patients with SGD, 11.2 mm (P = .003). There was also a significant difference

between patients with MGD and SGD in the number of lesions identified with 4D-CT (Table 2). A single candidate lesion was identified in 63% of patients with SGD on 4D-CT compared with 21% of those with MGD (P = .003). There was no statistically significant difference in mean calcium, PTH, and the Wisconsin Index between the 2 groups (Table 2).

Although notable differences were observed between MGD and SGD with regard to gland size and the number of lesions identified on 4D-CT, these factors individually did not serve as reliable predictors of MGD. Identification of either multiple or no abnormal glands was only 63% specific for MGD and 79% sensitive (Table 3). A lesion size of \leq 7 mm had a specificity of 79% for MGD with a sensitivity of 50%. A WIN of <800 had a 93% specificity for MGD but only a 25% sensitivity.

Performance of Composite MGD and 4D-CT MGD Scores

The mean composite MGD score was significantly higher among patients with MGD at 4.3 compared with those with SGD at 2.7 (P < .001). Similar findings were noted for the 4D-CT MGD score (2.9 versus 1.5, respectively; P < .001) (Table 2). Composite MGD scores of \geq 4, \geq 5, and 6 had high specificities of 72%, 86%, and 100%, respectively, for multigland disease (Table 4). When applied to the current prospective patient population, the calculated area under the receiver operating characteristic curve was 0.78 (Figure). The 4D-CT MGD scores of \geq 3 and 4 were similarly effective in predicting MGD, with specificities of 74% and 88%, respectively (Table 4), and with an area under the receiver operating characteristic curve of 0.78 (Fig 1*B*).

DISCUSSION

Preoperative localization of abnormal parathyroid glands plays an integral role in surgical planning. 4D-CT has improved localiza-

Table 3: Performance of individual components of MGD score system for predicting multigland disease

	No. of			Positive
	Patients	Sensitivity	Specificity	Predictive Value
No. of glands: ≥ 2 or no	22	79%	63%	58%
lesion on 4D-CT				
Size of gland: <7 mm	14	50%	79%	61%
or no lesion on 4D-CT				
WIN: <800	7	25%	93%	70%

Table 4: Performance of the composite and 4D-CT MGD scores for predicting multigland disease

	No. of Patients	Sensitivity	Specificity	Positive Predictive Value
	Fatients	Sensitivity	specificity	Fredictive value
Composite MGD score				
≥1	28	100%	5%	41%
≥2	26	93%	16%	42%
≥ 3	24	86%	51%	53%
≥ 4	22	79%	72%	65%
≥5	13	46%	86%	68%
6	6	21%	100%	100%
4D-CT MGD score				
≥ 0	28	100%	0%	39%
≥1	28	100%	26%	47%
≥2	23	82%	60%	58%
≥ 3	21	75%	74%	66%
4	9	32%	88%	64%

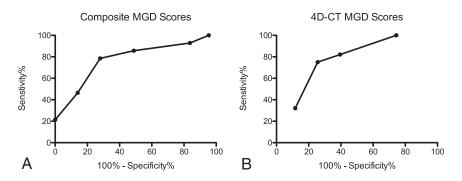


FIGURE. Receiver operating characteristic curves showing performance in predicting MGD with the composite MGD score (*A*) and the 4D-CT MGD score (*B*). The area under the curve is 0.78 for both methods.

tion ability in both MGD and SGD compared with ultrasound and sestamibi scans, but the sensitivity of 4D-CT for MGD remains low at 32%–53% compared with its sensitivity for SGD at 88%–93%.⁶⁻⁸ The composite MGD score and the 4D-CT MGD score systems proposed by Sepahdari et al¹¹ suggested that a multifactorial system using 4D-CT information, with or without laboratory information, could effectively categorize patients into subsets with a very high likelihood for MGD, a very high likelihood for SGD, or intermediate likelihood of either MGD or SGD. The major limitation of that study is that it was based on a retrospectively derived sample, without a separate validation cohort. In our current study, we applied the composite MGD and 4D-CT MGD score systems to a prospective cohort and confirmed that these score systems could be used to predict MGD with high specificity.

The retrospectively derived MGD score systems by Sepahdari et al¹¹ performed equally well in our prospective setting. Using a retrospective patient population, Sepahdari et al observed high specificities of 81%, 93%, and 98% in predicting multigland dis-

ease for composite MGD scores of ≥ 4 , \geq 5, and 6, respectively. In our current prospective study, similarly high specificities of 72%, 86%, and 100% were observed for composite MGD scores of \geq 4, \geq 5, and 6, respectively. This finding was also true for the 4D-CT score system. 4D-CT scores of \geq 3 and 4 yielded specificities of 81% and 96% in the retrospective population, compared with 74% and 88% in the current prospective population. Last, areas under the receiver operating characteristic curves of MGD predictive score systems were also similar between the retrospective and prospective patient cohorts (0.82 and 0.83 for composite and 4D-CT MGD scores in the retrospective cohort compared with 0.78 for both techniques in the prospective cohort). Our prospective cohort validated the retrospectively derived system of Sepahdari et al, achieving results nearly identical to those of the retrospective cohort.

Our study also shows that the number of candidate lesions identified on 4D-CT alone is not sufficient for distinguishing MGD from SGD. In clinical practice, radiologists may rely on detecting multiple glands as the best sign of MGD. However, we showed that identification of ≥ 2 lesions on 4D-CT is neither sensitive nor specific for MGD. Among patients with SGD, 12 of 43 (28%) were noted to have ≥ 2 suspected lesions (false-positive), and 6 of 28 (21%) patients with MGD had only 1 suspected lesion (false-negative). These findings illustrate the need for addi-

tional data regarding lesion size and biochemical information to better estimate the risk for MGD.

The composite and 4D-CT MGD score systems should serve as evidence-based foundations for communicating certainty and uncertainty in the diagnosis of abnormal parathyroid glands. They should maximize the clinically actionable information available from imaging. These scoring systems are valuable in giving objective probabilities of MGD versus SGD. In clinical practice, a radiologist can advise surgeons that a composite MDG score of \geq 5 or a 4D-CT MGD score of 4 is highly likely to represent MGD. Surgeons can then use these objective data to better plan operative approaches and counsel patients accordingly. For instance, given a high preoperative probability of MGD, surgeons may decide to avoid minimally invasive unilateral parathyroidectomy, which risks failing to identify abnormal glands during the operation, and instead opt for a 4-gland exploration to identify multiple abnormal parathyroid glands.

Several limitations exist in our study. Although this study was performed in a prospective manner, it remains a single-institution study with studies interpreted by a single neuroradiologist, who also participated in the original retrospective study. The composite MGD predictive score system is largely dependent on the 4D-CT findings, and imaging interpretation may vary in a systematic way among different radiologists. However, previous studies have shown high concordance in 4D-CT interpretations between neuroradiologists, suggesting that assignment of MGD scores is likely to be highly reproducible.^{14,15} Last, a patient population presenting to a single institution may represent inherent bias. Thus, the applicability of the MGD predictive score system should also be prospectively validated in a multi-institution setting. Notably, the previous study by Sepahdari et al,¹¹ in which the MGD predictive score system was developed, was performed in 2 academic institutions with 2 different interpreting neuroradiologists. Rates of lesion identification were similar between the 2 institutions and the MGD score system was developed on the basis of patients from both institutions.

CONCLUSIONS

The composite MGD score and 4D-CT MGD score performed well in a prospective setting, identifying patients with a high likelihood of MGD by using 4D-CT data and biochemical information. Use of these scores preoperatively to identify patients at risk for MGD may play an important role in operative planning.

REFERENCES

- 1. Sosa JA, Udelsman R. Minimally invasive parathyroidectomy. *Surg* Oncol 2003;12:125–34 CrossRef Medline
- Siperstein A, Berber E, Mackey R, et al. Prospective evaluation of sestamibi scan, ultrasonography, and rapid PTH to predict the success of limited exploration for sporadic primary hyperparathyroidism. Surgery 2004;136:872–80 CrossRef Medline
- McGreal G, Winter DC, Sookhai S, et al. Minimally invasive, radioguided surgery for primary hyperparathyroidism. Ann Surg Oncol 2001;8:856–60 CrossRef Medline
- Cayo AK, Sippel RS, Schaefer S, et al. Utility of intraoperative PTH for primary hyperparathyroidism due to multigland disease. Ann Surg Oncol 2009;16:3450–54 CrossRef Medline
- 5. Ruda JM, Hollenbeak CS, Stack BC Jr. A systematic review of the

diagnosis and treatment of primary hyperparathyroidism from 1995 to 2003. Otolaryngol Head Neck Surg 2005;132:359–72 CrossRef Medline

- 6. Rodgers SE, Hunter GJ, Hamberg LM, et al. **Improved preoperative** planning for directed parathyroidectomy with 4-dimensional computed tomography. *Surgery* 2006;140:932–40; discussion 940–41 CrossRef Medline
- Cheung K, Wang TS, Farrokhyar F, et al. A meta-analysis of preoperative localization techniques for patients with primary hyperparathyroidism. Ann Surg Oncol 2012;19:577–83 CrossRef Medline
- Kukar M, Platz TA, Schaffner TJ, et al. The use of modified fourdimensional computed tomography in patients with primary hyperparathyroidism: an argument for the abandonment of routine sestamibi single-positron emission computed tomography (SPECT). Ann Surg Oncol 2015;22:139–45 CrossRef Medline
- Harari A, Zarnegar R, Lee J, et al. Computed tomography can guide focused exploration in select patients with primary hyperparathyroidism and negative sestamibi scanning. *Surgery* 2008;144:970–76; discussion 976–79 CrossRef Medline
- Cham S, Sepahdari AR, Hall KE, et al. Dynamic parathyroid computed tomography (4DCT) facilitates reoperative parathyroidectomy and enables cure of missed hyperplasia. Ann Surg Oncol 2015; 22:3537–42 CrossRef Medline
- Sepahdari AR, Bahl M, Harari A, et al. Predictors of multigland disease in primary hyperparathyroidism: a scoring system with 4D-CT imaging and biochemical markers. *AJNR Am J Neuroradiol* 2015;36: 987–92 CrossRef Medline
- Mazeh H, Chen H, Leverson G, et al. Creation of a "Wisconsin index" nomogram to predict the likelihood of additional hyperfunctioning parathyroid glands during parathyroidectomy. Ann Surg 2013;257:138-41 CrossRef Medline
- 13. Siperstein A, Berber E, Barbosa GF, et al. **Predicting the success of limited exploration for primary hyperparathyroidism using ultrasound, sestamibi, and intraoperative parathyroid hormone: analysis of 1158 cases.** *Ann Surg* 2008;248:420–28 Medline
- Sepahdari AR, Yeh MW, Rodrigues D, et al. Three-phase parathyroid 4-dimensional computed tomography initial experience: inexperienced readers have high accuracy and high interobserver agreement. J Comput Assist Tomogr 2013;37:511–17 CrossRef Medline
- Bahl M, Sepahdari AR, Sosa JA, et al. Parathyroid adenomas and hyperplasia on four-dimensional CT scans: three patterns of enhancement relative to the thyroid gland justify a three-phase protocol. *Radiology* 2015;277:454–62 CrossRef Medline