

Natural History of Benign Solid and Cystic Thyroid Nodules

Erik K. Alexander, MD; Shelley Hurwitz, PhD; Jenny P. Heering, BA; Carol B. Benson, MD; Mary C. Frates, MD; Peter M. Doubilet, MD, PhD; Edmund S. Cibas, MD; P. Reed Larsen, MD; and Ellen Marqusee, MD

Background: Thyroid nodules are common and most often benign. The natural history of benign thyroid nodules, however, is unclear.

Objective: To determine the natural history of cytologically benign thyroid nodules using ultrasonography.

Design: Retrospective case series.

Setting: Single tertiary care clinic.

Participants: All patients referred to the Brigham and Women's Hospital Thyroid Nodule Clinic, Boston, Massachusetts, who had benign cytologic results on ultrasonography-guided fine-needle aspiration of a thyroid nodule between 1995 and 2000 and returned for a requested follow-up examination 1 month to 5 years later.

Measurements: Nodule dimensions were measured at both vis-

its, and growth was defined as an increase in calculated volume of 15% or greater. These results were correlated with the time between examinations, age, sex, baseline serum thyroid-stimulating hormone concentration, and cystic content of each nodule.

Results: Nodule volume increased over time ($P < 0.001$). The estimated proportion of nodules with an increase in volume of 15% or greater after 5 years was 89%. Nodules with greater cystic content were less likely to grow than solid nodules ($P = 0.01$). Seventy-four of the 330 nodules were reaspirated on the second visit. Despite an average increase in volume of 69%, only 1 of 74 reaspirated nodules was malignant.

Conclusion: Most solid, benign thyroid nodules grow. Therefore, an increase in nodule volume alone is not a reliable predictor of malignancy.

Ann Intern Med. 2003;138:315-318.

For author affiliations, see end of text.

www.annals.org

Thyroid nodules are present in nearly 50% of adults, increasing in prevalence with age (1). The evaluation of thyroid nodules that measure 1 cm or greater in diameter typically includes a screening measure of serum thyroid-stimulating hormone (TSH) levels and fine-needle aspiration (FNA). Most FNA results are benign (90% to 95%), and follow-up examinations are advised. Recommendations include periodic clinical examinations or ultrasonography, with or without suppressive L-thyroxine therapy (1, 2). Nodules that increase in size during follow-up are often regarded as suspicious for malignancy, and repeated FNA or surgery is advised (3–6). Data supporting these recommendations are limited, however, as few reports have evaluated thyroid nodule growth using the most sensitive technique, high-resolution ultrasonography. Furthermore, criteria defining nodule growth are inconsistent; some guidelines use an increase in maximal diameter of greater than 50%, while others suggest an increase in maximal diameter greater than 5 mm or an increase in calculated volume greater than 15% (5–10). We used ultrasonography of thyroid nodules to determine the natural history of cytologically benign thyroid nodules over a 1-month to 5-year follow-up period.

METHODS

We retrospectively reviewed the records of all patients referred to the dual-discipline Thyroid Nodule Clinic at Brigham and Women's Hospital, Boston, Massachusetts, for evaluation of nodular thyroid disease between 1995 and 2000. All patients referred to the clinic underwent ultrasonography of the thyroid by a radiologist and ultrasonography-guided FNA of nodules measuring 1 cm or greater in maximal diameter by an endocrinologist. All ultrasonog-

raphy evaluations were adequate for review and interpretation. All patients with benign cytology on initial FNA were advised to schedule follow-up ultrasonography 9 to 12 months later. Repeated FNA was performed on the follow-up visit at the discretion of the endocrinologist, usually because of nodule growth. The study sample included all patients with nodules with benign cytologic results on the initial visit who returned for follow-up ultrasonography within the 5-year period.

Thyroid ultrasonography was performed by one of three radiologists using a 5- to 15-MHz transducer. The length, width, and depth of each nodule were reported, and each nodule was classified as solid, less than 25% cystic, 25% to 50% cystic, 50% to 75% cystic, or greater than 75% cystic. Nodule volume was calculated by using the formula for a rotational ellipsoid ($\text{length} \times \text{width} \times \text{depth} \times \pi/6$) (7, 11, 12). Ultrasonography-guided FNA was performed with a 25-gauge needle (three to four aspirations per nodule), and specimens were processed by using the Thin-Prep technique (Cytec Corp., Boxborough, Massachusetts). All slides were read by a cytopathologist at Brigham and Women's Hospital. Specimens were considered benign when six or more groups of benign follicular cells (each group containing ≥ 15 cells) were identified without atypical features.

Repeated ultrasonographies were performed, and findings were directly compared with the previous images. Change in nodule size over the interval between examinations was assessed by using three criteria: 1) change in maximal diameter greater than 50% (7, 8, 12, 13); 2) change in maximal diameter of 3 mm or more; 3) change in calculated volume of 15% or more (11, 14). The latter

Context

Although benign thyroid nodules are common, we know relatively little about their natural history.

Contribution

This observational study from a single tertiary care facility used repeated ultrasonography to show that benign thyroid nodules typically increase in volume over a 3- to 5-year period. Solid nodules grew more than cystic nodules, and only 1 of 74 reaspirated nodules was malignant.

Implications

Nodule growth alone does not predict malignancy.

—The Editors

two criteria are defined by established inaccuracy rates for each method (11, 15).

The Investigational Review Board of Brigham and Women's Hospital granted permission to perform this review.

Descriptive statistics are presented according to nodule or patient as appropriate. The main outcome for the single-variable and multivariable predictive models was nodule growth, defined as an increase in volume of 15% or greater. Single-variable and multivariable mixed-effects logistic regression was used to predict growth, while accounting for the correlation structure in the data where some patients had more than one nodule (16). Potential predictors were the time between examinations, cystic content (solid, <25% cystic, 25% to 50% cystic, 50% to 75% cystic, or >75% cystic), TSH level (mIU/L), L-thyroxine use, age, and sex. Unadjusted and adjusted odds ratios and 95% CIs were calculated. Time to growth was determined by using life-table methods. Data were analyzed by using SAS software, version 8.2 (SAS Institute, Inc., Cary, North Carolina).

The funding sources had no role in the design, conduct, or reporting of the study or the decision to publish the manuscript.

RESULTS

A total of 1009 patients were examined in the Thyroid Nodule Clinic between 1995 and 2000, and 1358 nodules were biopsied. On initial FNA, 854 nodules (in 700 patients) measured 1 cm or greater in maximal diameter with benign cytologic results. Two hundred sixty-eight patients (38%) with 330 benign thyroid nodules (39%) returned for follow-up ultrasonography, with a mean interval of 20 months (range, 1 to 65 months) between examinations (Appendix Figure, available at www.annals.org). The baseline demographic and ultrasonography characteristics of these 268 patients and their nodules were similar to those of the 432 patients who did not return for follow-up (Table 1).

Change in nodule size over each patient's follow-up

period was assessed by three methods to facilitate comparison with previous studies. With use of a greater than 50% change in maximal diameter, 14 nodules (4%) were determined to have increased in size upon repeated ultrasonography. With evaluating change in maximal diameter of 3 mm or greater or change in calculated volume (cm³) of 15% or greater, 86 nodules (26%) and 129 nodules (39%), respectively, were determined to have increased in size on follow-up ultrasonography.

The time interval between examinations was significantly correlated with nodule growth ($r = 0.22$; $P < 0.001$). Table 2 shows mixed-models logistic regression analysis for prediction of thyroid nodule growth (volume change $\geq 15\%$). Time between examinations and lower cystic content remained statistically significant predictors of growth in the final multivariable model. Each year, the background odds of growth increased by 50%. The estimated median time to achieve volume growth of 15% or greater was 35 months (95% CI, 29 to 41 months). The estimated proportion with growth was 53% (CI, 46% to 61%) at 3 years and 89% (CI, 81% to 97%) at 5 years using life-table methods. The patient's age, sex, baseline serum TSH concentration, or L-thyroxine use did not predict nodule growth.

Sixty-one patients underwent repeated FNA at the time of the second ultrasonography. The nodules in this group were larger on initial examination (2.7 cm vs. 2.3 cm; $P = 0.001$) and had increased in volume by an average of 69% during follow-up compared with 14% in those nodules not rebiopsied ($P < 0.001$). Patient characteristics were similar except for a longer interval between examinations (28 months vs. 18 months) and younger age (43 years vs. 48 years) noted among the rebiopsied group. One of the 74 repeated FNA samples suggested a follicular neo-

Table 1. Demographic and Ultrasonography Characteristics of Patients with a Benign Thyroid Nodule ≥ 1 cm in Maximal Diameter Who Returned for Follow-up Ultrasonography as Recommended Compared with Those Who Did Not*

Characteristic	Patients without Follow-up Ultrasonography	Patients with Follow-up Ultrasonography	P Value
Patients, <i>n</i>	432	268	
Women, %	89	91	>0.2
Age, <i>y</i>	49 \pm 15	47 \pm 14	0.04
Thyroid-stimulating hormone level, <i>mIU/L</i>	1.45 \pm 1.27	1.4 \pm 0.9	0.19
Thyroid nodules, <i>n</i>	524	330	
Maximum diameter, <i>cm</i>	2.4 \pm 1.2	2.4 \pm 1.0	>0.2
Volume calculation, <i>cm</i> ³	6.4 \pm 11.0	5.5 \pm 9.1	0.14
Nodule characteristics, %			
Solid	45	39	>0.2
<25% cystic	24	27	
25%–50% cystic	11	13	
50%–75% cystic	8	10	
>75% cystic	12	11	

* Number of patients and thyroid nodules expressed as absolute numbers. Age, thyroid-stimulating hormone concentration, and nodule size are expressed as means \pm SD.

Table 2. Single-Variable Predictors and Final Multivariable Model To Predict Thyroid Nodule Growth (Volume Increase \geq 15%)

Variable	Odds Ratio (95% CI)*	P Value
Single-variable predictors		
Time between examinations		
1 mo	1.03 (1.02–1.05)	<0.001
6 mo	1.22 (1.11–1.35)	
12 mo	1.50 (1.24–1.81)	
Nodule characteristic		
>50% cystic	0.37 (0.19–0.69)	0.002
L-thyroxine use	0.81 (0.46–1.40)	>0.2
Thyroid-stimulating hormone level, mIU/L	0.87 (0.54–1.25)	>0.2
Age	1.00 (0.99–1.02)	>0.2
Male sex	1.99 (0.87–4.56)	0.10
Final multivariable model		
Time between examinations		
1 mo	1.03 (1.02–1.05)	<0.001
6 mo	1.21 (1.10–1.33)	
12 mo	1.47 (1.22–1.78)	
Nodule characteristic		
>50% cystic	0.43 (0.23–0.82)	0.01

* For single-variable predictors, odds ratios are unadjusted; for the final multivariable model, odds ratios are adjusted.

plasm, and the remainder were benign. The nodule was removed; it was a poorly differentiated papillary carcinoma. It had enlarged from 10.1 cm³ to 18.1 cm³ in volume (an 80% increase) over 38 months.

DISCUSSION

We used ultrasonography to assess the natural history of 330 benign thyroid nodules measuring 1 cm or more in maximal diameter followed for a mean period of 20 months. Although the 268 patients (with 330 nodules) represent only 39% of the benign nodules seen in the Brigham and Women's Hospital Thyroid Nodule Clinic between 1995 and 2000, they appear to be representative of the whole group with respect to demographic and nodule characteristics. Using the most rigorous criteria (\geq 15% increase in volume), we documented growth in 39% of benign thyroid nodules during follow-up, which indicates that many such nodules grow. Consistent with our findings, Brander and colleagues (10) found that 35% of benign nodules increased in size over 4.9 to 5.6 years. However, the criteria for growth were not defined, and minimal data on repeated FNA were provided (10). Similarly, Papini and colleagues (15) documented an increase in mean nodule volume among patients in the control group of a 5-year randomized study that assessed the efficacy of L-thyroxine suppression therapy for nodular goiter. Our results also support previous conclusions that more cystic nodules are less likely to grow compared with nodules with a greater solid component (5).

Current opinion suggests that increasing nodule size has modest but significant power for predicting thyroid cancer (2). Kuma and colleagues found malignancy in 26%

of previously unbiopsied nodules that increased in size over a 10- to 30-year period (5). A follow-up study 2 years later reported a malignancy rate of 4.5% among nodules that were previously found to be benign on FNA and subsequently grew (although no definition of growth was provided) (6). In our study, only 1 of 74 rebiopsied nodules was malignant on repeated FNA biopsy. Although only 22% of nodules seen in follow-up were rebiopsied, this group had statistically significantly larger nodules at baseline and grew more compared with those not rebiopsied (69% vs. 14% increase in volume; 2.9 mm vs. 0.1 mm increase in diameter). Thus, by growth criteria, these patients were at substantially higher risk. While the clinical history and rate of increase must be considered, our results indicate that most benign nodules grow slowly with time.

The differences in our conclusions from those of most earlier studies derive from the examination methods and the growth criteria. Several early studies used palpation to assess nodule size, which is a much less accurate method than ultrasonography (17). The definition of growth as an increase in maximal diameter of greater than 50% is based on such studies and is provided here to facilitate comparison of our results (7, 13). Only 4% of the nodules in our data set met this criterion, while 96% were unchanged, similar to the results of earlier studies.

We acknowledge potential limitations to our study. First, patients returning for follow-up examination were not chosen randomly, thereby allowing for potential physician (or patient) selection bias. However, at initial examination, all patients were instructed to arrange follow-up examinations in 9 to 12 months, regardless of patient or nodule characteristics. Table 1 confirms that the study cohort is similar to patients who did not return for follow-up ultrasonography. Second, the time when growth occurred cannot be known precisely and, thus, may overestimate the stated time estimates for nodule growth. Although this may affect the rate of nodule growth, however, it does not affect the overall conclusion that most benign nodules enlarge. Finally, we reaspirated only 74 of the 330 initially benign thyroid nodules. These patients had demographic characteristics similar to those of the nonrebiopsied group; however, their thyroid nodules increased in volume to a statistically significantly greater degree (69% vs. 14%), thus providing a rationale for repeated FNA. Earlier studies support the validity and accuracy of initial FNA assessment of nodules, especially those that are stable or smaller in size upon repeated examination (18, 19). Since all nodules were initially biopsied using ultrasonography guidance, we have a much lower concern for sampling error than we would if the initial aspirations were performed without assurance that the sampling site was within the nodule (20).

In summary, a large fraction of benign thyroid nodules grow over time. Thus, slow growth of a thyroid nodule proven benign by ultrasonography-guided FNA does not necessarily indicate a false-negative initial result or subse-

quent malignant transformation; rather, it represents the natural history of such nodules.

From Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts.

Acknowledgments: The authors thank Dr. Robert Utiger for his helpful review of and comments on this article.

Grant Support: By training grants from the National Institutes of Health (DK-07529 and HL-07609); research grant from The Endocrine Fellows Foundation; Thyroid Center of Excellence grant from Knoll Pharmaceuticals; and grants from Biostatistical Consulting Service, Center for Clinical Investigation, Brigham and Women's Hospital.

Requests for Single Reprints: Erik K. Alexander, MD, Endocrine Division, Brigham and Women's Hospital, 221 Longwood Avenue, 2nd Floor, Boston, MA 02115; e-mail, ekalexander@partners.org.

Current author addresses and contributions are available at www.annals.org.

References

- Mazzaferri EL. Management of a solitary thyroid nodule. *N Engl J Med*. 1993;328:553-9. [PMID: 8426623]
- Ridgway E. Clinical evaluation of solitary thyroid nodules. In: Braverman LE, Utiger RD, eds. *Werner & Ingbar's The Thyroid*. 8th ed. Philadelphia: Lippincott Williams & Williams; 2000:949-55.
- Singer PA, Cooper DS, Daniels GH, Ladenson PW, Greenspan FS, Levy EG, et al. Treatment guidelines for patients with thyroid nodules and well-differentiated thyroid cancer. American Thyroid Association. *Arch Intern Med*. 1996;156:2165-72. [PMID: 8885814]
- Daniels GH. Thyroid nodules and nodular thyroids: a clinical overview. *Compr Ther*. 1996;22:239-50. [PMID: 8733781]
- Kuma K, Matsuzuka F, Kobayashi A, Hirai K, Morita S, Miyachi A, et al. Outcome of long standing solitary thyroid nodules. *World J Surg*. 1992;16:583-7; discussion 587-8. [PMID: 1413828]
- Kuma K, Matsuzuka F, Yokozawa T, Miyachi A, Sugawara M. Fate of untreated benign thyroid nodules: results of long-term follow-up. *World J Surg*. 1994;18:495-8; discussion 499. [PMID: 7725734]
- Reverter JL, Lucas A, Salinas I, Audí L, Foz M, Sanmartí A. Suppressive therapy with levothyroxine for solitary thyroid nodules. *Clin Endocrinol (Oxf)*. 1992;36:25-8. [PMID: 1559296]
- Gharib H, James EM, Charboneau JW, Naessens JM, Offord KP, Gorman CA. Suppressive therapy with levothyroxine for solitary thyroid nodules. A double-blind controlled clinical study. *N Engl J Med*. 1987;317:70-5. [PMID: 3295553]
- Tannirandorn P, Fish SA, Langer JE, Mandel SJ. Benign thyroid nodules do not grow [Abstract]. Presented at the 83rd Meeting of the Endocrine Society, Denver, CO; 2001.
- Brander AE, Viikinkoski VP, Nickels JI, Kivisaari LM. Importance of thyroid abnormalities detected at US screening: a 5-year follow-up. *Radiology*. 2000;215:801-6. [PMID: 10831702]
- Brunn J, Block U, Ruf G, Bos I, Kunze WP, Scriba PC. [Volumetric analysis of thyroid lobes by real-time ultrasound (author's transl)]. *Dtsch Med Wochenschr*. 1981;106:1338-40. [PMID: 7274082]
- La Rosa GL, Lupo L, Giuffrida D, Gullo D, Vigneri R, Belfiore A. Levothyroxine and potassium iodide are both effective in treating benign solitary solid cold nodules of the thyroid. *Ann Intern Med*. 1995;122:1-8. [PMID: 7985890]
- Cheung PS, Lee JM, Boey JH. Thyroxine suppressive therapy of benign solitary thyroid nodules: a prospective randomized study. *World J Surg*. 1989;13:818-21; discussion 822. [PMID: 2696232]
- Berghout A, Wiersinga WM, Drexhage HA, Smits NJ, Touber JL. Comparison of placebo with L-thyroxine alone or with carbimazole for treatment of sporadic non-toxic goitre. *Lancet*. 1990;336:193-7. [PMID: 1973768]
- Papini E, Petrucci L, Guglielmi R, Panunzi C, Rinaldi R, Bacci V, et al. Long-term changes in nodular goiter: a 5-year prospective randomized trial of levothyroxine suppressive therapy for benign cold thyroid nodules. *J Clin Endocrinol Metab*. 1998;83:780-3. [PMID: 9506726]
- Liang KY. Longitudinal data analysis using generalized linear models. *Biometrika*. 1986;73:13-22.
- Brander A, Viikinkoski P, Tuuhea J, Voutilainen L, Kivisaari L. Clinical versus ultrasound examination of the thyroid gland in common clinical practice. *J Clin Ultrasound*. 1992;20:37-42. [PMID: 1309541]
- Erdoğan MF, Kamel N, Aras D, Akdoğan A, Başkal N, Erdoğan G. Value of re-aspirations in benign nodular thyroid disease. *Thyroid*. 1998;8:1087-90. [PMID: 9920362]
- Grant CS, Hay ID, Gough IR, McCarthy PM, Goellner JR. Long-term follow-up of patients with benign thyroid fine-needle aspiration cytologic diagnoses. *Surgery*. 1989;106:980-5; discussion 985-6. [PMID: 2588125]
- Danese D, Sciacchitano S, Farsetti A, Andreoli M, Pontecorvi A. Diagnostic accuracy of conventional versus sonography-guided fine-needle aspiration biopsy of thyroid nodules. *Thyroid*. 1998;8:15-21. [PMID: 9492148]

Current Author Addresses: Drs. Alexander, Hurwitz, Benson, Frates, Appendix Figure Overview of patients and thyroid nodules
Doubilet, Cibas, Larsen, and Marqusee and Ms. Heering: Endocrine included in analysis of nodule growth.
Division, Brigham and Women's Hospital, 221 Longwood Avenue, 2nd
Floor, Boston, MA 02115.

Author Contributions: Conception and design: E.K. Alexander, P.R.
Larsen, and E. Marqusee.

Analysis and interpretation of the data: E.K. Alexander, S. Hurwitz, J.P.
Heering, P.R. Larsen, and E. Marqusee.

Drafting of the article: E.K. Alexander, S. Hurwitz, J.P. Heering, E.S.
Cibas, P.R. Larsen, and E. Marqusee.

Critical revision of the article for important intellectual content: S. Hur-
witz, J.P. Heering, C.B. Benson, M.C. Frates, P.M. Doubilet, E.S. Ci-
bas, P.R. Larsen, and E. Marqusee.

Final approval of the article: E.K. Alexander, S. Hurwitz, M.C. Frates,
P.R. Larsen and E. Marqusee.

Provision of study materials or patients: P.R. Larsen and E. Marqusee.

Statistical expertise: S. Hurwitz.

Obtaining of funding: P.R. Larsen.

Administrative, technical, or logistic support: J.P. Heering, C.B. Benson,
P.M. Doubilet, E.S. Cibas, P.R. Larsen, and E. Marqusee.

Collection and assembly of data: E.K. Alexander, J.P. Heering, C.B.
Benson, M.C. Frates, P.M. Doubilet, and E. Marqusee.

FNA Fine-needle aspiration.