

Inter-Observer Variation in the Pathologic Identification of Minimal Extrathyroidal Extension in Papillary Thyroid Carcinoma

Henry K. Su,¹ Bruce M. Wenig,² Grace C. Haser,¹ Meghan E. Rowe,¹ Sylvia L. Asa,³ Zubair Baloch,⁴ Eugenie Du,⁵ William C. Faquin,⁶ Giovanni Fellegara,⁷ Thomas Giordano,⁸ Ronald Ghossein,⁹ Virginia A. LiVolsi,⁴ Ricardo Lloyd,¹⁰ Ozgur Mete,³ Umut Ozbek,¹¹ Juan Rosai,⁷ Saul Suster,¹² Lester D. Thompson,¹³ Andrew T. Turk,¹⁴ and Mark L. Urken¹⁵

Background: Extrathyroidal extension (ETE) is a significant prognostic factor in papillary thyroid carcinoma (PTC). Minimal extrathyroidal extension (mETE) is characterized by involvement of the sternothyroid muscle or perithyroid soft tissue, and is generally identified by light microscope examination. Patients with mETE, identified pathologically, are automatically upstaged to pT3. However, the prognostic implications of mETE have been a source of controversy in the literature. Moreover, there is also controversy surrounding the identification of mETE on pathological specimens. The objective of this study was to determine the level of agreement among expert pathologists in the identification of mETE in PTC cases.

Methods: Eleven expert pathologists from the United States, Italy, and Canada were asked to perform a review of 69 scanned slides of representative permanent sections of PTC specimens. Each slide was evaluated for the presence of mETE. The pathologists were also asked to list the criteria they use to identify mETE.

Results: The overall strength of agreement for identifying mETE was slight ($\kappa=0.14$). Inter-pathologist agreement was best for perithyroidal skeletal muscle involvement ($\kappa=0.46$, moderate agreement) and worst for invasion around thick-walled vascular structures ($\kappa=0.02$, slight agreement). In addition, there was disagreement over the constellation of histologic features that are diagnostic for mETE, which affected overall agreement for diagnosing mETE.

Conclusions: Overall agreement for the identification of mETE is poor. Disagreement is a result of both variation in individual pathologists' interpretations of specimens and disagreement on the histologic criteria for mETE. Thus, the utility of mETE in staging and treatment of PTC is brought into question. The lack of concordance may explain the apparent lack of agreement regarding the prognostic significance of this pathologic feature.

¹Department of Otolaryngology—Head and Neck Surgery, Thyroid, Head and Neck Cancer (THANC) Foundation, New York, New York.

Departments of ²Pathology and ¹⁵Otolaryngology—Head and Neck Surgery, Mount Sinai Beth Israel, New York, New York.

³Department of Pathology, Laboratory Medicine Program, Toronto General Hospital, Toronto, Canada.

⁴Department of Pathology and Laboratory Medicine, Hospital of the University of Pennsylvania, Philadelphia, Pennsylvania.

⁵Department of Otorhinolaryngology—Head and Neck Surgery, Montefiore Medical Center, Albert Einstein College of Medicine, Bronx, New York.

⁶Department of Pathology, Massachusetts General Hospital, Boston, Massachusetts.

⁷Centro Consulenze Anatomia Patologica Oncologica, Centro Diagnostico Italiano, Milan, Italy.

⁸Department of Pathology, University of Michigan Medical School, Ann Arbor, Michigan.

⁹Department of Pathology, Memorial Sloan-Kettering Cancer Center, New York, New York.

¹⁰Department of Pathology and Laboratory Medicine, University of Wisconsin, Madison, Wisconsin.

¹¹Population Health Science and Policy, Mount Sinai Hospital, New York, New York.

¹²Department of Pathology, Medical College of Wisconsin, Milwaukee, Wisconsin.

¹³Department of Pathology, Woodland Hills Medical Center, Woodland Hills, California.

¹⁴Department of Pathology, New York-Presbyterian/Columbia, New York, New York.

Introduction

EXTRATHYROIDAL EXTENSION (ETE) has long been cited as an important prognostic factor in the treatment of papillary thyroid cancer (PTC). The reported incidence of ETE in the literature varies between 5% and 45% (1). The following factors have shown to increase the risk of ETE: age ≥ 50 years, large tumor size (>4 cm), nonencapsulated tumors, aggressive histologic variants (e.g., tall-cell variant, diffuse sclerosing variant, hobnail variant), and the presence of nodal or distant metastasis (1). ETE increases the risk of both recurrence and mortality in patients with PTC (1–7). As such, the presence of ETE is used to upstage a patient with well-differentiated PTC in numerous staging systems (EORTC, AJCC-TNM, AMES, and MACIS) (8–11).

The seventh edition of the AJCC-TNM staging system for thyroid carcinomas divides ETE into minimal and significant ETE (12). Minimal ETE (mETE) is characterized by involvement of the sternothyroid muscle or perithyroid soft tissue, while significant ETE is characterized by invasion into subcutaneous soft tissue, larynx, trachea, esophagus, or recurrent laryngeal nerve. Significant ETE is usually appreciated at a macroscopic level in the operating room, whereas mETE is determined by light microscopic examination. Significant ETE can also be appreciated on ultrasound when capsular disruption is observed (13). Currently, mETE upstages a carcinoma measuring <4.0 cm to pT3, whereas significant ETE upstages a patient to pT4. While significant ETE remains an important factor in predicting prognosis, recurrence, and mortality in thyroid cancer cases, the impact of mETE on patient outcome is controversial. Several studies have found that mETE alone does not worsen survival rates in PTC (14,15). These findings raise the question of whether patients ought to be upstaged on the basis of mETE alone.

The authors recently reported the strong association between tumors with mETE and lymph node metastases showing extranodal extension (ENE). These co-occurrences may serve as an indicator of the biologic behavior of the primary tumor, which is conferred to the lymph node (16). ENE has been well documented to affect prognosis poorly (17). The correlation between mETE and ENE suggests that there is a reason to identify mETE, and that the presence of mETE should upstage the primary tumor. The biology of a tumor that displays ENE and/or mETE may be more aggressive and warrant treatment that is more vigorous.

There is controversy as to what histologic criteria constitute mETE. Mete *et al.* reviewed the anatomy of the neck organs, including the thyroid gland, and emphasized that the thyroid gland does not have a defined, limiting fibrous capsule but rather an incomplete, thin fibroadipose pseudocapsule. As such, Mete *et al.* proposed that involvement of adipose or fibrous tissue does not necessarily indicate ETE of the tumor (18). Furthermore, though skeletal muscle invasion does represent extrathyroidal spread, this is only reliable in the lateral lobes of the thyroid, and may not be reliable in the isthmus, since fibers of Soemmerring's muscle (a small strap muscle arising from the hyoid bone and inserting partly on the thyroid cartilage and partly on the isthmus, also known as *musculus levator glandulae thyroidae* of Soemmerring) can be identified in this particular location (18). Further, mesenchymal tissues normally located outside the thyroid gland, including fibrous and adipose tissue, nerve, muscle, and thick-walled blood vessels, can be

found within the substance of the thyroid gland. Such findings may complicate the evaluation of mETE, making the actual pathologic diagnosis of mETE subjective and problematic (18).

Controversy surrounding the prognostic significance of mETE may be related to disagreement among pathologists about the definition of mETE and discordance in its identification. To the authors' knowledge, there are no reported studies assessing the degree of concordance among pathologists in the assessment of mETE. This study evaluates inter-observer variation in the histologic identification of mETE. The importance of this analysis is designed to alert clinicians to any variability that may exist and also to determine whether a lack of concordance may be a root cause for the controversy regarding the prognostic significance of this histologic feature.

Methods

Institutional Review Board approval was obtained at Mount Sinai Beth Israel for this retrospective review. Eleven expert pathologists from the United States, Italy, and Canada were asked to perform a review of 69 scanned slides of representative permanent sections of PTC specimens from the Mount Sinai Beth Israel Medical Center in New York. Cases were selected on the basis of availability in the department at the time the study was undertaken. One slide was chosen because of limitations of digital scanning available from a non-departmental (commercial) vendor. The best slide in a given case was chosen for scanning. Before scanning, slides were selected specifically for being potentially controversial with respect to the presence of mETE, as determined by the lead pathologist (B.M.W.). All slides were prepared with hematoxylin and eosin stain in accordance with standard procedures at the Mount Sinai Beth Israel Medical Center. The slides were digitally scanned at 40 \times using the Aperio ScanScope AT Turbo. All 11 pathologists evaluated all cases.

Pathologists reviewed each slide and designated it as demonstrating mETE or having no evidence of mETE, specifically evaluating the specimen for perithyroidal involvement of fat, skeletal muscle, nerves, and thick-walled vascular structures. Involvement of thick-walled vascular structures included growth of tumor cells/nests around and beyond thick-walled vascular blood vessels located at the periphery of the sections. As involvement of the recurrent laryngeal nerve would be quantified as significant ETE, pathologists were aware in advance that any nerve involvement did not include the recurrent laryngeal nerve. The specific criteria used by each pathologist for diagnosing mETE were recorded. Each pathologist was also asked to explain if he or she used different criteria to classify mETE for a primary cancer arising from the isthmus. The pathologists involved in the study were informed that this study would be to evaluate digitally scanned slides in cases of PTC for the presence or absence of minimal mETE in order to evaluate the concordance or lack thereof among expert thyroid pathologists.

The kappa coefficient was computed to assess consistency in the pathologists' categorization of each slide and the criteria used. All statistical analyses were done with R statistical package v3.1.1.

Results

The kappa coefficient for the presence of mETE was 0.14, suggesting that the overall strength of agreement in identifying

TABLE 1. KAPPA COEFFICIENT FOR INTER-PATHOLOGIST AGREEMENT OF PRESENCE OF mETE AND INVOLVEMENT OF PERITHYROIDAL STRUCTURES

Criteria for mETE	Kappa coefficient	Strength of agreement ^a	Proportion of observed agreement
Presence of mETE	0.14	Slight	0.59
Perithyroidal fat involvement	0.16	Slight	0.58
Perithyroidal nerve involvement	0.11	Slight	0.88
Perithyroidal skeletal muscle involvement	0.46	Moderate	0.91
Perithyroidal thick-walled vessel involvement	0.02	Slight	0.67

^aKappa agreement: <0, less than chance agreement; 0.00–0.20, slight agreement; 0.21–0.40, fair agreement; 0.41–0.60, moderate agreement; 0.61–0.80, substantial agreement; 0.81–0.99, almost perfect agreement.

mETE, minimal extrathyroidal extension.

mETE was slight (Table 1). Inter-pathologist agreement was best for perithyroidal skeletal muscle involvement ($\kappa=0.46$, moderate agreement) and worst for involvement of thick-walled vascular structures ($\kappa=0.02$, slight agreement).

Seven out of 11 pathologists reported using perithyroidal fat, skeletal muscle, nerve, and thick-walled vessel (Fig. 1) involvement as histologic criteria for diagnosing mETE (Table 2). Two pathologists reported using all of these criteria except perithyroidal thick-walled vessel involvement for reporting mETE, and two pathologists reported using only

perithyroidal skeletal muscle involvement for identifying mETE. Five pathologists also reported using desmoplasia as a criterion for diagnosing mETE. Additional criteria noted include “lack of fat” surrounding the tumor and “extrinsic extension into vessels, not tumor emboli.”

The data were analyzed with data removed from two pathologist raters who only considered perithyroidal skeletal muscle involvement as evidence of mETE (pathologists 6 and 9; Table 2) because their criteria for mETE were quite different compared with the other nine pathologists. When the data from pathologists 6 and 9 were removed from the sampling, the kappa values increased slightly (Table 3). The kappa coefficient for the overall presence of mETE increased from 0.14 to 0.23. Kappa values also increased slightly for perithyroidal fat and thick-walled vessel involvement. However, kappa values decreased slightly for perithyroidal nerve and skeletal muscle involvement. While there was a small increase in kappa for some values, this was not statistically significant.

Criteria used for diagnosing mETE in primary tumors arising from the isthmus of the thyroid varied as well. Five pathologists stated they do not use different criteria to diagnose mETE in the isthmus, while six reported that they use more stringent criteria. Of these six, four reported that they consider skeletal muscle invasion near the isthmus to be less important, and two reported that adipose tissue involvement near the isthmus is less important.

Discussion

The pathologic diagnosis of mETE is subjective, without clearly established histopathologic diagnostic criteria, and the overall inter-pathologist concordance among 11 expert pathologists was poor. There was variation in individual pathologists' interpretations of the presence of mETE, as well

FIG. 1. Papillary thyroid carcinoma with (A) involvement of perithyroidal fat (hematoxylin and eosin [H&E]; 40 \times); (B) perineural invasion (arrowheads)—note the thick-walled blood vessel to the extreme right (H&E; 20 \times); (C) invasion into and beyond skeletal muscle (H&E; 10 \times); and (D) extension around perithyroidal thick-walled blood vessels (arrows) (H&E; 10 \times).

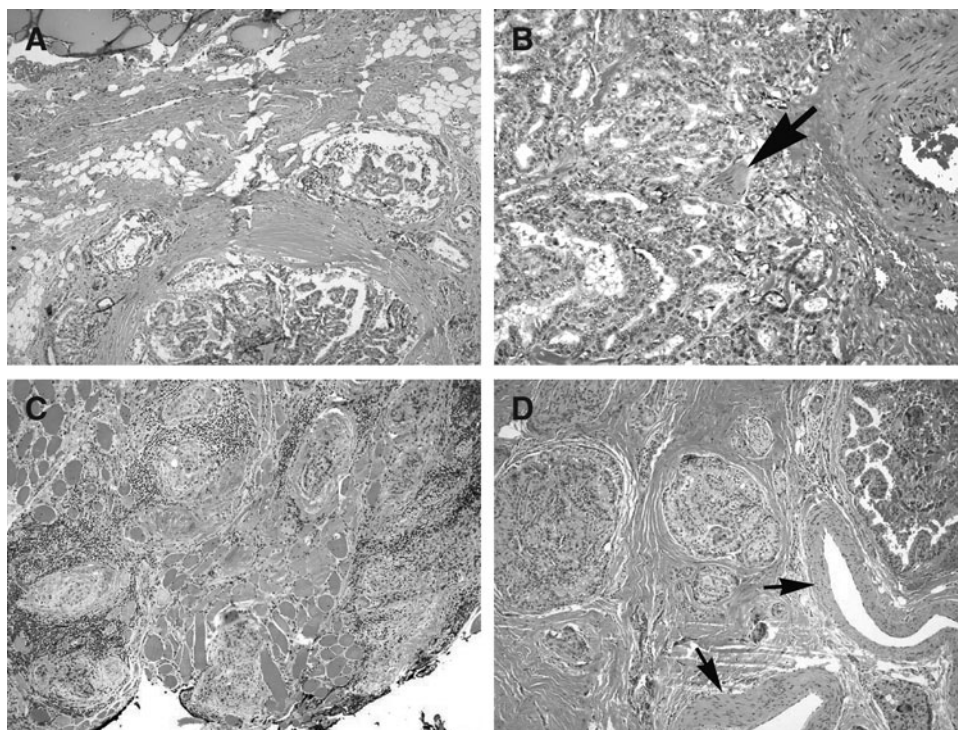


TABLE 2. HISTOLOGIC CRITERIA UTILIZED BY PATHOLOGISTS FOR DIAGNOSIS OF METE

	<i>Pathologist</i>										
	1	2	3	4	5	6	7	8	9	10	11
Histologic criteria for mETE utilized											
Perithyroidal fat involvement	Y	Y	Y	Y	Y	N	Y	Y	N	Y	Y
Perithyroidal skeletal muscle involvement	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Perithyroidal nerve involvement	Y	Y	Y	Y	Y	N	Y	Y	N	Y	Y
Perithyroidal thick-walled vessel involvement	Y	Y	Y	Y	Y	N	Y	N	N	N	Y
Other:											
Invasion of tumor rel. to normal thyroid		Desmo-plastic/reactive response; lack of fat	Desmo-plastic/reactive response; lack of fat	Fibrosis w/desmoplasia at thyroid periph.	Plane of thyroid edge	N/A	Desmo-plastic stromal reaction; plane of thyroid edge	Extrinsic extension into vessels, not tumor emboli	N/A	N/A	Desmo-plastic stromal reaction
Less emphasis on tumor in adipose tissue; more stringent criteria	Less emphasis on tumor in adipose tissue; more stringent criteria	None	None	Less emphasis on tumor in adipose tissue	None	Skeletal muscle invasion not evidence of mETE	Less emphasis on skeletal muscle	Muscle involvement must be extra-thyroidal	Skeletal muscle invasion not evidence of mETE	None	None

TABLE 3. KAPPA COEFFICIENT FOR INTER-PATHOLOGIST AGREEMENT OF PRESENCE OF mETE AND INVOLVEMENT OF PERITHYROIDAL STRUCTURES, NINE RATERS

<i>Criteria for mETE</i>	<i>Kappa coefficient</i>	<i>Strength of agreement^a</i>	<i>Proportion of observed agreement</i>
Presence of mETE	0.23	Fair	0.69
Perithyroidal fat involvement	0.20	Slight/fair	0.61
Perithyroidal nerve involvement	0.10	Slight	0.86
Perithyroidal skeletal muscle involvement	0.41	Moderate	0.90
Perithyroidal thick-walled vessel involvement	0.03	Slight	0.63

^aKappa agreement: <0, less than chance agreement; 0.00–0.20, slight agreement; 0.21–0.40, fair agreement; 0.41–0.60, moderate agreement; 0.61–0.80, substantial agreement; 0.81–0.99, almost perfect agreement.

as whether individual tissues (fat, muscle, nerve and thick-walled vessel) were affected. Additionally, there was disagreement in the criteria used by individual pathologists to identify mETE. When the findings of two pathologists with very different definitions of mETE (pathologists 6 and 9) were removed from the statistical analysis, kappa coefficients did not change significantly.

Controversy as to what histologic findings constitute mETE appears to stem partially from disagreement over whether perithyroidal invasion of fat and nerves around or beyond thick-walled vascular structures constitutes evidence for mETE (18). Since thyroid tissue may be present as a normal histologic variant within perithyroidal fat and skeletal muscle, and, alternatively, fat and muscle may be found within the thyroid (especially at the periphery of the gland, including the isthmus), identifying involvement of these types of tissue does not necessarily indicate the extrathyroidal extent of the tumor (18). One pathologist specified that only extrinsic extension into blood vessels, rather than tumor emboli, would constitute mETE.

The kappa coefficients were low for the identification of all features and for the determination of mETE, even when outlier pathologists were removed from the analysis. However, these kappa values are affected by the low number of events in the data (i.e., positive identification of mETE). In other words, raters overall identified involvement of perithyroidal structures and mETE rarely in the 69 slides studied. Therefore, it is important to place the kappa values in the context of the proportion of agreement. The proportion of agreement was >0.50 for all values, indicating that the kappa coefficients would have been higher had there been a more representative sample, with more cases of mETE.

The proportion of agreement was high for perithyroidal nerve and skeletal muscle involvement (0.88 and 0.91, respectively) compared with the proportion of agreement for the presence of mETE overall (0.59). This finding suggests that a part of the disagreement between pathologists on the presence of mETE lies not in the identification of tumor in-

vasion in various perithyroidal tissues, but in the decision to interpret a constellation of such histologic features as diagnostic for mETE. Therefore, clear delineation of the histologic criteria for mETE may have an important role to play in improving uniformity in the diagnosis of mETE. However, based on low kappa coefficients and only moderate proportions of agreement for the identification of perithyroidal fat and thick-walled vessel involvement, there appears to be disagreement in identifying these features, which also contributes to the low kappa coefficient for the overall presence of mETE. Further evaluation of this finding is warranted with a larger sample. Additionally, a study is warranted that would examine the prognostic implications of these variations in the determination of mETE.

There are several potential limitations of this study. First, the sample size included only 11 pathologists. Second, senior pathologists were utilized from various medical centers in the United States, Canada, and Italy who are considered experts in thyroid pathology. This naturally biases the group toward pathologists who are more experienced with mETE and PTC. This group is not necessarily representative of all pathologists who review thyroid specimens less frequently. It is possible that the degree of concordance for diagnosing mETE would be lower for all pathologists than for pathologists with extensive expertise in this area. However, given the low kappa values for even these senior pathologists, it is unclear if this would be the case. A second study designed to engage pathologists in community hospitals would help elucidate this point. A third limitation is that the pathologists were only sent one representative slide from each specimen. In practice, pathologists often review multiple slides and have the ability to probe further should a particular diagnosis be in question. A fourth limitation is that the 69 cases used in this study were selected because they were determined to be potentially controversial for the presence of mETE by the senior pathologist; this could have influenced the kappa values and resulted in lower agreement than in non-controversial cases. Moreover, since the 11 pathologists knew the cases were being selected for their complexity and ambiguity, this could have biased their interpretation. A fifth limitation is that the pathologists could have been affected by the knowledge that they were graded based on their interpretation.

mETE is a controversial prognostic factor in PTC. From this study, it is apparent that inter-observer concordance for the identification of mETE between expert pathologists is moderate to poor. Hence, thyroid cancer patients may be incorrectly identified as having mETE, or mETE may be missed in some patients. Recognition of this variation in pathologic interpretation is important for clinicians to understand when making clinical decisions regarding their patients and when reviewing the literature related to this feature. The results of this study support the need for standardization of histopathologic criteria for mETE to achieve more consistency among pathologists from different institutions.

Conclusions

This study evaluated inter-observer concordance in the histologic identification of mETE and proves that the overall agreement for the identification of mETE is poor. Disagreement is a result of both variation in individual pathologists'

interpretations and disagreement on the histologic criteria for mETE. Thus, the reliability of pathologists to be consistent in their interpretation and definition of mETE in PTC is questioned. Moreover, given the strong association between mETE and ENE, correct identification of mETE has prognostic implications. Clinicians must be aware of this inherent variability in reporting mETE. This may confound the interpretation of outcomes studies, which heavily rely upon the uniform reporting of pathologic features, and this impacts the staging of thyroid cancer patients.

Acknowledgments

The authors would like to acknowledge the Mount Sinai Health System for its generous support of this research project.

Author Disclosure Statement

The authors have nothing to disclose.

References

- Ortiz S, Rodriguez J, Soria T, Perez-Flores D, Pinero A, Moreno J, Parrilla P 2001 Extrathyroid spread in papillary carcinoma of the thyroid: clinicopathological and prognostic study. *Otolaryngol Head Neck Surg* **124**:261–265.
- Andersen PE, Kinsella J, Loree TR, Shaha AR, Shah JP 1995 Differentiated carcinoma of the thyroid with extrathyroidal extension. *Am J Surg* **170**:467–470.
- McConahey WM, Hay ID, Woolner LB, van Heerden J, Taylor WF 1986 Papillary thyroid cancer treated at the Mayo Clinic, 1946 through 1970: initial manifestations, pathologic findings, therapy, and outcome. *Mayo Clin Proc* **61**:978–996.
- Carcangiu ML, Zampi G, Pupi A, Castagnoli A, Rosai J 1985 Papillary carcinoma of the thyroid. *Cancer* **55**:805–828.
- Cody III HS, Shah JP 1981 Locally invasive, well-differentiated thyroid cancer: 22 years' experience at Memorial Sloan–Kettering Cancer Center. *Am J Surg* **142**:480–483.
- McCaffrey TV, Bergstralh EJ, Hay ID 1994 Locally invasive papillary thyroid carcinoma: 1940–1990. *Head Neck* **16**:165–172.
- Cushing SL, Palme CE, Audet N, Eski S, Walfish PG, Freeman JL 2004 Prognostic factors in well-differentiated thyroid carcinoma. *Laryngoscope* **114**:2110–2115.
- Byar DP, Green SB, Dor P, Williams ED, Colon J, van Gilse HA, Mayer M, Sylvester RJ, Van Glabbeke M 1979 A prognostic index for thyroid carcinoma. A study of the EORTC Thyroid Cancer Cooperative Group. *Eur J Cancer* **15**:1033–1041.
- Greene FL, Page DL, Fleming ID, Fntz AG, Balch CM, Haller DG, Morrow M (Eds) 2002 AJCC Cancer Staging Manual. Vol. 1. Springer, New York.
- Cady B, Rossi R 1988 An expanded view of risk-group definition in differentiated thyroid carcinoma. *Surgery* **104**:947–953.
- Hay ID, Bergstralh EJ, Goellner JR, Ebersold J, Grant C 1993 Predicting outcome in papillary thyroid carcinoma: development of a reliable prognostic scoring system in a cohort of 1779 patients surgically treated at one institution during 1940 through 1989. *Surgery* **114**:1050–1057.
- Edge SB, Byrd DR, Compton CC, Fritz AG, Greene FL, Trotti A (Eds) 2010 AJCC Cancer Staging Manual. Vol. 649. Springer, New York.
- Kwak JY, Kim E-K, Youk JH, Kim MJ, Son EJ, Choi SH, Oh KK 2008 Extrathyroid extension of well-differentiated papillary thyroid microcarcinoma on US. *Thyroid* **18**:609–614.
- Ito Y, Tomoda C, Uruno T, Takamura Y, Miya A, Kobayashi K, Matsuzuka F, Kuma K, Miyauchi A 2006 Minimal extrathyroid extension does not affect the relapse-free survival of patients with papillary thyroid carcinoma measuring 4 cm or less over the age of 45 years. *Surg Today* **36**:12–18.
- Shin JH, Ha TK, Park HK, Ahn MS, Kim KH, Bae KB, Kim TH, Choi CS, Kim TK, Bae SK 2013 Implication of minimal extrathyroidal extension as a prognostic factor in papillary thyroid carcinoma. *Int J Surg* **11**:944–947.
- Clain JB, Scherl S, Dos Reis L, Turk A, Wenig BM, Mehra S, Karle WE, Urken ML 2014 Extrathyroidal extension predicts extranodal extension in patients with positive lymph nodes: an important association that may affect clinical management. *Thyroid* **24**:951–957.
- Randolph GW, Duh Q-Y, Heller KS, LiVolsi VA, Mandel SJ, Steward DL, Tufano RP, Tuttle RM 2012 The prognostic significance of nodal metastases from papillary thyroid carcinoma can be stratified based on the size and number of metastatic lymph nodes, as well as the presence of extranodal extension. *Thyroid* **22**:1144–1152.
- Mete O, Rotstein L, Asa SL 2010 Controversies in thyroid pathology: thyroid capsule invasion and extrathyroidal extension. *Ann Surg Oncol* **17**:386–391.

Address correspondence to:

Grace C. Haser, BA
 Thyroid, Head and Neck Cancer (THANC) Foundation
 10 Union Square East, Suite 5B
 New York, NY 10003

E-mail: ghaser@thancfoundation.org