Thyroid Ultrasound Reporting Lexicon: White Paper of the ACR Thyroid Imaging, Reporting and Data System (TIRADS) Committee

Edward G. Grant, MD\textsuperscript{a}, Franklin N. Tessler, MD\textsuperscript{b}, Jenny K. Huang, MBBS\textsuperscript{c}, Jill E. Langer, MD\textsuperscript{d}, Michael D. Beland, MD\textsuperscript{d}, Lincoln L. Berland, MD\textsuperscript{e}, John J. Cronan, MD\textsuperscript{f}, Terry S. Desser, MD\textsuperscript{f}, Mary C. Frates, MD\textsuperscript{g}, Ulrike M. Hamper, MD\textsuperscript{h}, William D. Middleton, MD\textsuperscript{i}, Carl C. Reading, MD\textsuperscript{j}, Leslie M. Scoutt, MD\textsuperscript{k}, A. Thomas Stavros, MD\textsuperscript{l}, Sharlene A. Teefey, MD\textsuperscript{l}

Abstract

Ultrasound is the most commonly used imaging technique for the evaluation of thyroid nodules. Sonographic findings are often not specific, and definitive diagnosis is usually made through fine-needle aspiration biopsy or even surgery. In reviewing the literature, terms used to describe nodules are often poorly defined and inconsistently applied. Several authors have recently described a standardized risk stratification system called the Thyroid Imaging, Reporting and Data System (TIRADS), modeled on the BI-RADS system for breast imaging. However, most of these TIRADS classifications have come from individual institutions, and none has been widely adopted in the United States. Under the auspices of the ACR, a committee was organized to develop TIRADS. The eventual goal is to provide practitioners with evidence-based recommendations for the management of thyroid nodules on the basis of a set of well-defined sonographic features or terms that can be applied to every lesion. Terms were chosen on the basis of demonstration of consistency with regard to performance in the diagnosis of thyroid cancer or, conversely, classifying a nodule as benign and avoiding follow-up. The initial portion of this project was aimed at standardizing the diagnostic approach to thyroid nodules with regard to terminology through the development of a lexicon. This white paper describes the consensus process and the resultant lexicon.

Key Words: Thyroid nodule, ultrasound, thyroid cancer, structured reporting, thyroid imaging

INTRODUCTION

The incidence of thyroid nodules has increased tremendously in recent years. The reasons for this increase are likely multifactorial but are largely attributed to widespread application of high-resolution ultrasound to the thyroid itself and the frequent incidental detection of nodules on other imaging modalities. In distinction to palpation, which demonstrates nodules in only 5% to 10% of the population, autopsy and sonography detect them in at least 60% [1]. Although nodules are extremely common, the incidence of malignancy in them is relatively low, ranging between 1.6% and 12% [2,3].
Ultrasound is superior to other modalities in characterizing thyroid nodules. Unfortunately, the findings are often not specific, and definitive diagnosis usually requires fine-needle aspiration (FNA) biopsy or even surgery. Because nodules are so common, a significant burden is placed on the health care system, and considerable anxiety may occur in patients. Furthermore, the majority of thyroid cancers are of the papillary type, which is typically indolent. Long-term studies by Ito et al [4] showed no difference in outcomes between patients with biopsy-proven carcinomas <1 cm undergoing thyroidectomy and those followed with no surgical intervention.

The literature regarding thyroid nodule characterization with ultrasound is expansive, and several professional organizations have put forth position or consensus statements. The two best known in the United States are from the American Thyroid Association (ATA) and the Society of Radiologists in Ultrasound [5,6]. Because FNA biopsy is such an integral part of the workup of thyroid nodules, the American Society of Cytopathology convened its own consensus panel to standardize reporting of biopsy results, which is known as the Bethesda classification [7].

Several authors have suggested a standardized risk stratification system called the Thyroid Imaging, Reporting and Data System (TIRADS), modeled on the BI-RADS® system for breast imaging, which has received widespread acceptance [8-10]. These proposals include the initial report by Horvath et al [9], as well as subsequent proposals by Kwak et al [8] and Park et al [10]. Despite these efforts, none of these TIRADS classifications have been widely adopted, particularly in the United States.

Our objective, therefore, was to develop a practical, standard lexicon for describing the sonographic characteristics of thyroid nodules, with the ultimate aim of applying it to risk stratification and triage of nodules for consistent follow-up in clinical practice.

METHODS

Beginning in 2012, a group of radiologists with expertise in thyroid imaging undertook a three-stage process under the auspices of the ACR; a subcommittee was charged with completing each one. The first effort, led by Lincoln Berland, MD, and Jenny Hoang, MBBS, was aimed at proposing recommendations for nodules discovered incidentally on imaging. This work led to a white paper published in 2015 [11]. The work reported here on the development of an ultrasound lexicon was led by Edward Grant, MD, whereas the final stage, which will be directed at risk stratification on the basis of the lexicon, will be led by Franklin Tessler, MD.

After an extensive literature review, relevant articles were distributed to all subcommittee members. Each radiologist was assigned three or four articles and was asked to list terms used by the authors to describe thyroid nodules. ACR staff members collated the lists, and a master list was drawn up. Frequency of use was the initial guide for determining which terms would be included in the lexicon.

The committee initially identified nine categories or families of terms that could be applied to all thyroid nodules: nodule composition, echogenicity, characteristics of cystic/solid components, shape, size/dimensions, margins, halo, echogenic foci, and flow/Doppler. Next, subcommittee members re-reviewed the literature to determine whether there was evidence that the categories and terms had discriminatory value in distinguishing benign from malignant nodules, which led to culling the category list. This process resulted in the selection of six final categories. Several of the original categories as well as numerous terms were eliminated from the lexicon or incorporated into other groups based either on infrequency of use or lack of statistical agreement with regard to their diagnostic value. Two members were assigned to develop definitions for each category and its individual terms in a format used for other ACR “RADS” lexicons.

THYROID ULTRASOUND CATEGORIES

Category 1: Composition

Definition.

- Composition describes the internal components of a nodule, that is, the presence of soft tissue or fluid, and the proportion of each.
  - Solid: Composed entirely or nearly entirely of soft tissue, with only a few tiny cystic spaces (Fig. 1A).
  - Predominately solid: Composed of soft tissue components occupying 50% or more of the volume of the nodule (Fig. 1B, online only).
  - Predominately cystic: Composed of soft tissue components occupying less than 50% of the volume of the nodule (Fig. 1C, online only).
  - Cystic: Entirely fluid filled.
Spongiform: Composed predominately of tiny cystic spaces (Fig. 1D, online only).

Background and Significance.
- A nodule should fit into one of the foregoing five categories. However, rarely, it may be difficult to determine if a nodule is filled with hemorrhagic material or is solid. Color Doppler flow may be useful in differentiating between the two.
- Papillary thyroid carcinoma (PTC) is most commonly solid, but many solid nodules are also benign; a solid nodule has a 15% to 27% chance of being malignant [6]. Some nodules undergo cystic degeneration or necrosis. A recent study of partially cystic nodules showed that the prevalence of malignancy was low whether the nodule was predominately cystic (6.1%) or predominately solid (5.7%) [12]. When evaluating a partially cystic nodule, it is important to evaluate the solid component. If the solid component is eccentrically (peripherally) located within a partially cystic nodule and the margin of the solid component has an acute angle with the wall of the nodule, the risk for malignancy is increased. Furthermore, if the solid component is hypoechoic, is lobulated, has an irregular border or punctate echogenic foci line, or has vascular flow, the risk for malignancy is increased. If the solid component is isoechoic, is centrally located within the nodule or, if peripheral, has no acute angle with the nodule wall, or has a smooth margin, spongiform appearance, or comet tail artifacts, it is likely benign [13,14].
- Purely cystic nodules or spongiform nodules have a very low risk for malignancy [6]. Two definitions of spongiform nodules have been proposed in the literature. When a spongiform nodule was defined as “the aggregation of multiple microcystic components in more than 50% of the volume of the nodule,” only one in 52 spongiform nodules was malignant [15]. When a spongiform nodule was defined as tiny cystic spaces involving the entire nodule, all 210 spongiform nodules were benign on FNA biopsy [16].

Comment.
- Terms used to describe partially cystic or partially solid nodules vary, with some authors breaking down the ratio of the two into numerical values on the basis of percentage whereas others choose to be more descriptive. The committee believed that the simple, subjective description of predominately cystic versus predominately solid would suffice.
- Although somewhat controversial, the committee agreed that finding several tiny cystic spaces in an otherwise completely solid nodule would still allow it to be classified as solid.

Category 2: Echogenicity

Definition.
- Level of echogenicity of the solid, noncalcified component of a nodule, relative to surrounding thyroid tissue [8,10,15,17-19].
  - Hyperechoic: Increased echogenicity relative to thyroid tissue (Fig. 2A, online only).
  - Isoechoic: Similar echogenicity relative to thyroid tissue.
  - Hypoechoic: Decreased echogenicity relative to thyroid tissue (Fig. 2B, online only).
  - Very hypoechoic: Decreased echogenicity relative to adjacent neck musculature (Fig. 2C).

Background and Significance.
- The level of nodule echogenicity is associated with both benign and malignant lesions. Very hypoechoic nodules have low sensitivity but very high specificity. Hypo-echogenicity is more sensitive but does not have high specificity [15,18,19].

Comment.
- The echogenicity of the solid component of a nodule should be compared with normal-appearing thyroid tissue, usually immediately adjacent to the nodule. In the setting of background abnormal thyroid tissue echogenicity, such as in Hashimoto’s thyroiditis, the echogenicity of the solid component should still be
described relative to the adjacent thyroid tissue, but it may be noted that the background tissue is of altered echogenicity.

- If the nodule is of mixed echogenicity, it can be described as “predominantly” hyperechoic, isoechoic, or hypoechoic.

**Category 3: Shape**

**Term:** taller-than-wide.

**Definition.**

- A taller-than-wide shape is defined as a ratio of >1 in the anteroposterior diameter to the horizontal diameter when measured in the transverse plane (Fig. 3).

**Background and Significance.**

- Taller-than-wide shape is a major feature for the categorization of thyroid nodules that are suspicious or suggestive of malignancy. The corresponding pathologic feature leading to this appearance is thought to be decreased compressibility. This finding is seen in 12% of thyroid nodules [8]. Sensitivity ranges between 40% and 68%, specificity between 82% and 93%, positive predictive value between 0.58 and 0.73, and negative predictive value between 0.77 and 0.88 [8,17,20-22].

**Comment.**

- In studies that specify how measurements are made, there are no significant differences comparing transverse or longitudinal dimensions [20,23]. For simplicity and consistency, the committee chose the ratio of >1 in the anteroposterior diameter to the horizontal diameter in the transverse plane.

**Category 4: Size**

**How the nodule should be measured:**

- Use maximal diameter on the basis of longitudinal, anteroposterior, and transverse measurements in centimeters per millimeter.

**Background and Significance.**

- Multiple studies have suggested that nodule size is not an independent predictor of malignancy risk in PTC. Tiny nodules can harbor malignancy, and large nodules are often benign. In a Finnish autopsy study of 101 thyroid glands, investigators found small, occult thyroid cancers in 36% [24]. As noted previously, the study by Ito et al [4] showed no value in performing thyroidectomy on small cancers.

- The correlation between nodule size and risk for malignancy remains controversial for nodules >1 cm. A 2013 study that included 7,346 nodules examined the effect of nodule size on the prevalence of thyroid cancer. At a threshold of 2 cm, the investigators found a statistically significant increase in cancer rate: 10.5% among the nodules 1 to 1.9 cm in diameter versus 15% for nodules >2 cm. Whether this is clinically significant is doubtful. In this study, larger nodules, when cancerous, were significantly more likely to have a histology other than papillary carcinoma (follicular, Hurthle cell, or other rare malignancies) [25].
Comment.

- Current thinking about biopsy of nodules <1 cm seems to be shifting to a more conservative approach. The most recent ATA guidelines [26] do not recommend biopsy of most lesions <1 cm. For nodules larger than 1 cm, considering the uncertainty between nodule size and malignancy risk, compared with the more consistent data on the impact of other sonographic features, we believe it is reasonable not to include size in the TIRADS scoring system.

Category 5: Margins

Definition.

- Refers to the border or interface between the nodule and the adjacent thyroid parenchyma or adjacent extrathyroidal structures.
  
  ○ Smooth: Uninterrupted, well-defined, curvilinear edge typically forming a spherical or elliptical shape (Fig. 4A, online only)
  
  ○ Irregular margin: The outer border of the nodule is spiculated, jagged, or with sharp angles with or without clear soft tissue protrusions into the parenchyma. The protrusions may vary in size and conspicuity and may be present in only one portion of the nodule (Fig. 4B).
  
  ○ Lobulated: Border has focal rounded soft tissue protrusions that extend into the adjacent parenchyma. The lobulations may be single or multiple and may vary in conspicuity and size (small lobulations are referred to as microlobulated) (Fig. 4C, online only).
  
  ○ Ill-defined: Border of the nodule is difficult to distinguish from thyroid parenchyma; the nodule lacks irregular or lobulated margins.
  
  ○ Halo: Border consists of a dark rim around the periphery of the nodule. The halo can be described as completely or partially encircling the nodule. In the literature, halos have been further characterized as uniformly thin, uniformly thick, or irregular in thickness.
  
  ○ Extrathyroidal extension: Nodule extends through the thyroid capsule (Fig. 4D, online only).

Background and Significance.

- A smooth border is more common in benign nodules, but between 33% and 93% of malignancies may have smooth borders [15,27]. Irregular and lobulated margins are features suspicious for thyroid malignancy [28]. These borders are considered to represent an aggressive growth pattern, although regions of thyroiditis can also have irregular margins. An ill-defined thyroid nodule margin has not been shown to be statically significantly associated with malignancy and is a common finding in benign hyperplastic nodules and thyroiditis [15,27,29].

- A halo may be due to a true fibrous capsule or a pseudocapsule. A uniform halo suggests a benign nodule because most thyroid malignancies are unencapsulated. However, a complete or incomplete halo has been noted in 10% to 24% of thyroid carcinomas.

- Extension of the nodule through the thyroid capsule into the adjacent soft tissue structures indicates invasive malignancy [29].

Comment.

- Analysis of the literature about the reported sensitivity and specificity of margin features is challenging because of previous nonuniformity in the definitions and the high rate of interobserver variability [15].

Category 6: Echogenic Foci

Definition.

- Refers to focal regions of markedly increased echogenicity within a nodule relative to the surrounding tissue. Echogenic foci vary in size and shape and may be encountered alone or in association with several well-known posterior acoustic artifacts.
  
  ○ Punctate echogenic foci: “Dot-like” foci having no posterior acoustic posterior artifacts. Kwak et al [8] defined punctate foci/microcalcifications as being <1 mm. Most authors define this feature on the basis of appearance alone (Fig. 5A).

Fig 4B. Irregular margin: 47-year-old woman with heterogeneously hyperechoic 16-mm nodule with irregular margins. Note angulated borders anteriorly. Diagnosis: papillary carcinoma. Figures 4A, 4C, 4D available online.
Macrocalcifications: When calcifications become large enough to result in posterior acoustic shadowing, they should be considered macrocalcifications. Macrocalcifications may be irregular in shape (Fig. 5B, online only).

Peripheral calcifications: These calcifications occupy the periphery of the nodule. The calcification may not be completely continuous but generally involves the majority of the margin. Peripheral calcifications are often dense enough to obscure the central components of the nodule (Fig. 5B, online only).

Comet-tail artifacts: A comet-tail artifact is a type of reverberation artifact. The deeper echoes become attenuated and are displayed as decreased width, resulting in a triangular shape. If an echogenic focus does not have this feature, a comet-tail artifact should not be described (Fig. 5D, online only).

Background and Significance.

- Echogenic foci have been associated with both benign and malignant lesions.
- Many authors have referred to all punctate echogenic foci in thyroid nodules simply as microcalcifications, but the majority of such foci are found in benign nodules, and therefore the term microcalcification is a misnomer [12]. The origins of punctate echogenic foci other than from true psammomatous microcalcifications of PTC are likely varied but, for example, have been shown to arise from the back walls of tiny unresolved cysts when seen in other organs such as the ovary or kidney. Although seen in both benign and malignant nodules, multiple studies have shown high specificity for punctate echogenic foci in malignant nodules [15,19,30,31].
- Macrocalcifications are generally considered to have an association with increased risk for malignancy, perhaps slightly more than double the baseline risk [8,15,21].
- For peripheral calcifications, studies have been conflicting, with some demonstrating an increased association with malignancy and others not [12,22].
- Recently, authors have subclassified comet-tail artifacts into small and large types and found a prevalence of malignancy of 15% in nodules that had echogenic foci with small comet-tail artifacts [12]. Conversely, when considering large comet-tail artifacts in cystic or partially cystic nodules, multiple studies have shown a strong association with benignity [12,32,33].

Comment.

- When present, the type of echogenic focus encountered within a given nodule should be specified. If more than one type of echogenic focus is present in a single nodule, each should be enumerated. If none are present, this should be stated.

DISCUSSION

Ultrasound is the most commonly used imaging technique in the evaluation of thyroid nodules, and its use has increased the discovery of nodules greatly. With that in mind, and given that there are conflicting recommendations from several societies, the ACR sponsored this TIRADS project. In keeping with other similar projects, the eventual goal of TIRADS is to provide practitioners with evidence-based recommendations formulated upon defined sonographic features of a given nodule. This initial portion of our project was aimed at standardizing the diagnostic approach to thyroid nodules with regard to terminology.

A wide array of terms has been used to describe the characteristics of thyroid nodules. Closely examining the existing literature, we found that terms used to describe nodules are often poorly defined and inconsistently applied. Furthermore, multiple terms have often been used to describe the same feature. This has led to confusion as to when and how these terms should be applied and, in many cases, what they actually mean. Clearly, inconsistent reporting leads to confusion about recommendations for further management.
Our committee sought to use terms already in common use in the literature rather than inventing new ones. We sought to provide concise written definitions with illustrations that can be used as a guide for practitioners. Terms that were chosen demonstrated consistency with regard to performance in the diagnosis of thyroid cancer or to classifying a nodule as benign. Additionally, it was believed that features should be included that would be reproducible among various readers.

Although Doppler was considered as a potential category, the committee did not recommend its inclusion on the basis of inconsistent literature about its value in differentiating cancer from benign nodules. One might apply it technically (eg, to better define a subtle nodule), but this is not something that would influence the TIRADS classification.

Another problematic subject was nodule size. Again, there is no consistent information in the literature that equates increasing size to an increased incidence of cancer. On the other side of the size spectrum, the recent guidelines published by the ATA [26] state that only nodules greater than 1 cm should be evaluated, unless there are compelling circumstances, such as a lesion that bulges the thyroid capsule or has accompanying abnormal lymph nodes or nodules in patients with high-risk clinical factors. Reviewing the literature, we found that characteristics of thyroid nodules other than size are more closely associated with benign or malignant lesions, and we prefer that these form the basis for our standardized evaluation of thyroid nodules.

### TAKE-HOME POINTS

- The goal of this project is to standardize terminology applied to thyroid nodules, such that all practitioners approach their evaluation in a similar fashion.
- The establishment of a lexicon is an essential initial step that provides a structured method for evaluation. The imager is directed to evaluate specific aspects of the nodule in an orderly fashion and is provided with a set of well-defined terms and features within each area of that evaluation.
- The aim is to decrease the variation seen in reporting of thyroid nodules in current practice. The ultimate goal is to develop guidelines for follow-up on the basis of statistics associated with the terms in the lexicon.

### REFERENCES

20. Moon HJ, Kwak JY, Kim EK, Kim MJ. A taller-than-wide shape in thyroid nodules in transverse and longitudinal ultrasonographic
planes and the prediction of malignancy. Thyroid 2011;21:1249-53.


33. Beland MD, Kwon L, Delellis RA, Cronan JJ, Grant EG. Non-shadowing echogenic foci in thyroid nodules are certain appearances enough to avoid thyroid biopsy? J Ultrasound Med 2011;30:753-60.
ACR Thyroid Imaging, Reporting and Data System (TI-RADS): White Paper of the ACR TI-RADS Committee

Franklin N. Tessler, MD, CM, William D. Middleton, MD, Edward G. Grant, MD, Jenny K. Hoang, MBBS, Lincoln L. Berland, MD, Sharlene A. Teefey, MD, John J. Cronan, MD, Michael D. Beland, MD, Terry S. Deser, MD, Mary C. Frates, MD, Lynwood W. Hammers, DO, Ulrike M. Hamper, MD, Jill E. Langer, MD, Carl C. Reading, MD, Leslie M. Scoutt, MD, A. Thomas Stavros, MD

Abstract
Thyroid nodules are a frequent finding on neck sonography. Most nodules are benign; therefore, many nodules are biopsied to identify the small number that are malignant or require surgery for a definitive diagnosis. Since 2009, many professional societies and investigators have proposed ultrasound-based risk stratification systems to identify nodules that warrant biopsy or sonographic follow-up. Because some of these systems were founded on the BI-RADS® classification that is widely used in breast imaging, their authors chose to apply the acronym TI-RADS, for Thyroid Imaging, Reporting and Data System. In 2012, the ACR convened committees to (1) provide recommendations for reporting incidental thyroid nodules, (2) develop a set of standard terms (lexicon) for ultrasound reporting, and (3) propose a TI-RADS on the basis of the lexicon. The committees published the results of the first two efforts in 2015. In this article, the authors present the ACR TI-RADS Committee’s recommendations, which provide guidance regarding management of thyroid nodules on the basis of their ultrasound appearance. The authors also describe the committee’s future directions.

Key Words: Thyroid nodule, thyroid cancer, management guidelines, ultrasound

INTRODUCTION
Thyroid nodules are exceedingly common, with a reported prevalence of up to 68% in adults on high-resolution ultrasound [1]. Currently, fine-needle aspiration (FNA) is the most effective, practical test to determine whether a nodule is malignant or may require surgery to reach a definitive diagnosis [2]. However, most nodules are benign, and even malignant nodules, particularly ones...
smaller than 1 cm, frequently exhibit indolent or nonaggressive behavior [3-5]. Therefore, not all detected nodules require FNA and/or surgery.

Despite a rapid increase in the reported incidence of papillary thyroid cancer that resulted from screening thyroid sonography in asymptomatic patients in South Korea, mortality has remained extremely low [6]. In the United States, overdiagnosis of thyroid cancer, defined as “diagnosis of thyroid tumors that would not, if left alone, result in symptoms or death” accounted for 70% to 80% of thyroid cancer cases in women and 45% of cases in men between 2003 and 2007 [7].

Therefore, a reliable, noninvasive method to identify which nodules warrant FNA on the basis of a reasonable likelihood of biologically significant malignancy would be highly desirable. In 2015, committees convened by the ACR published white papers that presented an approach to incidental thyroid nodules and proposed standard terminology (lexicon) for ultrasound reporting [8,9]. The purpose of the present white paper is to present our system for risk stratification, which is designed to identify most clinically significant malignancies while reducing the number of biopsies performed on benign nodules.

PROJECT RATIONALE AND CONSENSUS PROCESS

Several professional societies and groups of investigators have proposed methods to guide ultrasound practitioners in recommending FNA on the basis of ultrasound features [10-18]. Some of these systems were termed TI-RADS (Thyroid Imaging, Reporting and Data System) because they were modeled on the ACR’s BI-RADS®, which has been widely accepted in breast imaging. Other societies, such as the American Thyroid Association (ATA), have taken a slightly different, pattern-oriented approach, but with the same intent [19]. The plethora, complexity, and lack of congruence of these systems has limited their adoption by the ultrasound community and inspired our effort to publish a classification system under the auspices of the ACR. The ACR TI-RADS Committee agreed on the following attributes for our risk classification algorithm. It would be:

- founded on the ultrasound features defined in our previously published lexicon;
- easy to apply across a wide gamut of ultrasound practices;
- able to classify all thyroid nodules; and
- evidence based to the greatest extent possible.

The proposals presented in this white paper, which were developed via conference calls, e-mail, and online surveys, represent the consensus opinion of the ACR TI-RADS Committee. They are based on the literature; analysis of data from the Surveillance, Epidemiology, and End Results (SEER) Program of the National Cancer Institute; evaluation of existing risk classification systems; and expert opinion. Our recommendations are intended to serve as guidance for practitioners who incorporate ultrasound in the management of adult patients with thyroid nodules. They should not be construed as standards. Interpreting and referring physicians are legally and professionally responsible for applying their professional judgment to every case, regardless of the ACR TI-RADS recommendations. The decision to perform FNA should also account for the referring physician’s preference and the patient’s risk factors for thyroid cancer, anxiety, comorbidities, life expectancy, and other relevant considerations.

OVERVIEW OF ACR TI-RADS

The ultrasound features in the ACR TI-RADS are categorized as benign, minimally suspicious, moderately suspicious, or highly suspicious for malignancy. Points are given for all the ultrasound features in a nodule, with more suspicious features being awarded additional points. Figure 1 presents these features arranged per the five lexicon categories [8]. When assessing a nodule, the reader selects one feature from each of the first four categories and all the features that apply from the final category and sums the points. The point total determines the nodule’s ACR TI-RADS level, which ranges from TR1 (benign) to TR5 (high suspicion of malignancy). Note that although it is possible for a nodule to be awarded zero points and hence be characterized as TR1, all other nodules merit at least two points because a nodule that has a mixed cystic and solid composition (one point) will also gain at least one more point for the echogenicity of its solid component. Finally, although sonoelastography is a promising technique [20,21], it is probably not available in many ultrasound laboratories and is not incorporated into the ACR TI-RADS.

In the ACR TI-RADS, recommendations for FNA or ultrasound follow-up are based on a nodule’s ACR TI-RADS level and its maximum diameter. For risk levels TR3 through TR5, the chart presents a size threshold at or above which FNA should be recommended. We also defined lower size limits for recommending follow-up ultrasound for TR3, TR4, and TR5 nodules to limit the number of repeat sonograms for those that are likely to be benign or not clinically significant.
In developing the ACR TI-RADS, the committee strived to account for the discrepancy between the sharp rise in the diagnosis and treatment of thyroid cancer resulting from increased detection and biopsy and the lack of commensurate improvement in long-term outcomes [22]. This suggested to us that diagnosing every thyroid malignancy should not be our goal. Like other professional societies [17,19], we recommend biopsy of thyroid malignancy should not be our goal. Like other professional societies [17,19], we recommend biopsy of thyroid malignancy should not be our goal.

The ACR TI-RADS is designed to balance the benefit of identifying clinically important cancers against the risk and cost of subjecting patients with benign nodules or indolent cancers to biopsy and treatment. Our recommendations for follow-up ultrasound substantially mitigate the possibility that significant malignancies will remain undetected over time and are concordant with the increasing trend toward active surveillance (“watchful waiting”) for low-risk thyroid cancer [23,24].

DIFFERENCES BETWEEN ACR TI-RADS AND OTHER SYSTEMS

Structure
To make the system easy to understand and apply, the ACR TI-RADS does not include subcategories, nor does it include a TR0 category to indicate a normal thyroid gland. The ACR TI-RADS also lends itself to implementation as templates in voice recognition reporting or computerized decision support systems. The committee decided against the pattern-based approach used by the ATA on the basis of the results of a study by Yoon et al [25], which showed that the ATA guidelines were unable to classify 3.4% of 1,293 nodules, of which 18.2% were malignant. Notably, that study included only nodules that were subjected to FNA or surgery. It is likely that...
an even greater percentage of nodules would not have been categorized had other nodules been included, as it is not practical to provide patterns that account for every potential constellation of features.

**Size Thresholds for FNA**
The ACR TI-RADS is consistent with most other guidelines in recommending FNA for highly suspicious nodules 1 cm or larger. However, our thresholds for mildly suspicious and moderately suspicious nodules (2.5 and 1.5 cm, respectively) are higher than the cutoffs advocated by the ATA and the Korean Society of Thyroid Radiology [17,19].

In a 2005 publication cited by both organizations, Machens et al [26] contended that the cumulative risk for distant metastases from papillary and follicular thyroid cancer rose at a threshold of 2 cm. They therefore advocated biopsy of nodules larger than 2 cm. However, our review of their graphs suggested a gradual, slight increase that began at a larger size. More important, Machens et al based their analysis on tumor size in resected specimens, not on ultrasound dimensions. Subsequent research has demonstrated a significant lack of concordance between sonographic and pathologic sizing, with a tendency for ultrasound to result in larger measurements [27]. Of 205 papillary carcinomas \( \geq 1.5 \) cm studied by Bachar et al [28], the mean diameter on ultrasound was 2.65 \( \pm 1.07 \) cm, compared with 1.97 \( \pm 1.17 \) cm on pathology. The committee’s higher size cutoffs reflect this discrepancy.

Our choices were further guided by evaluation of a database of more than 3,000 proven thyroid nodules that was created for a study sponsored by the Society of Radiologists in Ultrasound [29]. Partial analysis showed cancer risk levels of no more than 2\% for TR1 and TR2 nodules, 5\% for TR3 nodules, 5\% to 20\% for TR4 nodules, and at least 20\% for TR5 nodules. We also considered published [30] and newly performed SEER data analyses that showed a slight increase in distant metastases at 2.5 cm, as well as slight increments in 10-year relative and thyroid cancer-specific mortality at 3 cm.

**ACR TI-RADS FEATURE CATEGORIES**
In this section, we elaborate on the five groups of ultrasound findings, ACR TI-RADS levels, and size thresholds. Readers are encouraged to refer to the lexicon white paper for detailed descriptions of all the categories and features [8]. As well, any history of prior FNA or ethanol ablation should be sought, as these procedures may lead to a suspicious appearance at follow-up ultrasound [31].

**Composition**
Nodules that are cystic or almost completely cystic merit no points because they are almost universally benign [16]. Similarly, a spongiform architecture is highly correlated with benign cytology, regardless of its relative echogenicity or other features [32-34]. However, a spongiform nodule must be composed predominantly (>50\%) of small cystic spaces [8]. Nodules should not be characterized as spongiform solely on the basis of the presence of a few, scattered cystic components in an otherwise solid nodule (Online Fig. 2).

“Mixed cystic and solid” combines two features from the lexicon, predominately solid and predominately cystic. The appearance of the solid component is more important than the overall size of the nodule or the proportion of solid versus cystic components in determining whether biopsy is warranted. Solid material that is eccentric and has an acute angle with the nodule’s wall is suspicious, as is solid material with moderately or highly suspicious characteristics, such as decreased echogenicity, lobulation, and punctate echogenic foci [35-38]. As well, although color Doppler ultrasound has not been shown to reliably discriminate between benign and malignant nodules [39], the presence of flow in solid components distinguishes tissue from echogenic debris or hemorrhage. Inconsequential debris may be identified by layering or motion elicited by changes in patient position.

**Echogenicity**
This feature refers to a nodule’s reflectivity relative to adjacent thyroid tissue, except for very hypoechoic nodules, in which the strap muscles are used as the basis for comparison. This category also includes “anechoic,” a zero-point feature that was absent from the lexicon. It applies to cystic or almost completely cystic nodules that would otherwise be given three points because of their very hypoechoic appearance.

**Shape**
A taller-than-wide shape is an insensitive but highly specific indicator of malignancy [12,13,40]. This feature is evaluated in the axial plane by comparing the height (“tallness”) and width of a nodule measured parallel and perpendicular to the ultrasound beam, respectively. A taller-than-wide configuration is usually evident on visual inspection and rarely requires formal measurements.

**Margin**
The presence of a halo is neither discriminatory nor mutually exclusive with other margin types; therefore, we
elected to omit it. We included “ill defined” in this group so that any reporting template that incorporates a field for margin will not be left empty if a nodule is not well defined. “Lobulated or irregular margin” refers to a spiculated or jagged edge, with or without protrusions into the surrounding parenchyma. It may be difficult to recognize this finding if the nodule is ill defined, is embedded in a heterogeneous gland, or abuts multiple other nodules. If the margin cannot be determined for any reason, zero points should be assigned.

Extension beyond the thyroid border is classified as extensive or minimal [41]. (We use the term border because the thyroid gland lacks a true fibrous capsule [42].) Extensive extrathyroidal extension (ETE) that is characterized by frank invasion of adjacent soft tissue and/or vascular structures is a highly reliable sign of malignancy and is an unfavorable prognostic sign [43]. Minimal ETE may be suspected sonographically in the presence of border abutment, contour bulging, or loss of the echogenic thyroid border [44,45]. However, agreement among pathologists for identification of minimal ETE is poor [46], and its clinical significance is controversial [41,47,48]. Therefore, practitioners should exercise caution when reporting minimal ETE, particularly for otherwise benign-appearing nodules.

Echogenic Foci
“Large comet-tail artifacts” are echogenic foci with V-shaped echoes >1 mm deep to them. They are associated with colloid and are strongly indicative of benignity when found within the cystic components of thyroid nodules. “Macrocalcifications” are coarse echogenic foci accompanied by acoustic shadowing. Evidence in the literature regarding their association with increased malignancy risk is mixed, especially in nodules lacking other malignant features [12,49-51]. Given published data that show a weakly positive relationship with malignancy [52], macrocalcifications are assigned one point, recognizing that the risk is increased if the nodule also contains moderately or highly suspicious features that warrant additional points.

Peripheral calcifications lie along all or part of a nodule’s margin. Their correlation with malignancy in the literature is variable [49]. However, because some publications suggest that they are more strongly associated with malignancy than macrocalcifications, they are awarded two points [52]. Some authors have called attention to interrupted peripheral calcifications with protruding soft tissue as suspicious for malignancy, but with low specificity [53]. In the ACR TI-RADS, this appearance qualifies as a lobulated margin, which adds another two points to the nodule’s total assignment.

In nodules with calcifications that cause strong acoustic shadowing that precludes or limits assessment of internal characteristics, particularly echogenicity and composition, it is best to assume that the nodule is solid and assign two points for composition and one point for echogenicity.

Punctate echogenic foci are smaller than macrocalcifications and are nonshadowing. In the solid components of thyroid nodules, they may correspond to the psammomatous calcifications associated with papillary cancers and are therefore considered highly suspicious, particularly in combination with other suspicious features. This category includes echogenic foci that are associated with small comet-tail artifacts in solid components, as distinguished from the large comet-tail artifacts described earlier [54]. Notably, small echogenic foci may be seen in spongiform nodules, where they probably represent the back walls of minute cysts. They are not suspicious in this circumstance and should not add to the point total of spongiform nodules.

Additional BenignAppearances
Several ultrasound findings have been described as characteristic of benign nodules with a high degree of reliability. These include a uniformly hyperechoic (“white knight”) appearance, as well as a variegated pattern of hyperechoic areas separated by hypoechoic bands reminiscent of giraffe hide, both in the setting of Hashimoto’s thyroiditis [34]. Because of their scarcity, the committee chose not to formally incorporate these patterns in the ACR TI-RADS chart.

PAPILLARY THYROID MICROCARCINOMAS
The ACR TI-RADS is concordant with other guidelines in recommending against routine biopsy of nodules smaller than 1 cm, even if they are highly suspicious. However, because some thyroid specialists advocate active surveillance, ablation, or lobectomy for papillary microcarcinomas, biopsy of 5- to 9-mm TR5 nodules may be appropriate under certain circumstances [24,55-57]. The determination to perform FNA will involve shared decision making between the referring physician and the patient. The report should indicate whether the nodule can be measured reproducibly on follow-up studies. Additionally, nodules in critical submarginal locations may complicate surgery [24,55,56]. Therefore, the report should also indicate whether the nodule abuts the trachea or whether it is adjacent to the tracheoesophageal groove (the location of the recurrent laryngeal nerve).
REPORTING CONSIDERATIONS

Measurement and Documentation

Accurate sizing of thyroid nodules is critical, as the maximum dimension determines whether a given lesion should be biopsied or followed. Although some interobserver discrepancy is inevitable because of variable conspicuity, consistent technique improves measurement accuracy and reproducibility.

Nodules should be measured in three axes: (1) maximum dimension on an axial image, (2) maximum dimension perpendicular to the previous measurement on the same image, and (3) maximum longitudinal dimension on a sagittal image (Online Fig. 3) [58]. For obliquely oriented nodules, these measurements may be different than the ones used to determine a taller-than-wide shape, but this discrepancy should rarely present a problem in practice. Measurements should also include the nodule’s halo, if present. Practitioners may use linear dimensions to determine volume, a calculation that is available on many ultrasound machines. Regardless of the method used, each practice should measure and report nodules consistently to facilitate serial comparison.

Meticulous documentation of the location of nodules on sonograms is equally important, particularly when the thyroid gland is heterogeneous or multiple nodules are present. Every nodule targeted for surveillance should be numbered sequentially and labeled as to its location in the thyroid gland (right, left, isthmus, upper, mid, lower, and, if necessary, lateral, medial, anterior, or posterior). Ultrasound video clips are valuable to provide further information about the spatial relationships between nodules and adjacent structures. The committee recommends that no more than four nodules with the highest ACR TI-RADS point scores that fall below the size threshold for FNA should be followed, as detailed reporting of more than four nodules would needlessly complicate and lengthen reports. Other nodules may be reassessed on subsequent sonograms without being formally enumerated.

The committee did not address follow-up for previously sampled nodules. The decision to request a repeat biopsy is typically driven by referring physicians and is guided by management recommendations based on prior FNA results based on the Bethesda System for Reporting Thyroid Cytopathology [19,59,60].

Definition of Growth

Criteria for significant growth depend on the size of the nodule and must take measurement variability into account. In the ACR TI-RADS, significant enlargement is defined as a 20% increase in at least two nodule dimensions and a minimal increase of 2 mm, or a 50% or greater increase in volume, as in the criteria adopted by other professional societies [19]. Because enlargement may not be apparent if the current sonogram is compared only with the immediately preceding one, it is important to also review measurements from earlier scans, if available.

Timing of Follow-Up Sonograms

There is little consensus in the literature regarding optimal spacing of follow-up sonograms for nodules that do not meet size criteria for FNA, as growth rates do not reliably distinguish benign from malignant nodules [61]. The committee believes that scanning intervals of less than 1 year are not warranted [62], except for proven cancers under active surveillance, which may require more frequent follow-up at the discretion of the referring physician [56]. We advocate timing on the basis of a nodule’s ACR TI-RADS level, with additional sonograms for lesions that are more suspicious. For a TR5 lesion, we recommend scans every year for up to 5 years. For a TR4 lesion, scans should be done at 1, 2, 3, and 5 years. For a TR3 lesion, follow-up imaging may be performed at 1, 3, and 5 years. Imaging can stop at 5 years if there is no change in size, as stability over that time span reliably indicates that a nodule has a benign behavior [63]. There is no published evidence to guide management of nodules that enlarge significantly but remain below the FNA size threshold for their ACR TI-RADS level at 5 years, but continued follow-up is probably warranted. If a nodule’s ACR TI-RADS level increases on follow-up, the next sonogram should be done in 1 year, regardless of its initial level.

Number of Nodules to Biopsy

Biopsy of three or more nodules is poorly tolerated by patients and increases cost with little or no benefit and some added risk. Therefore, the committee recommends targeting no more than two nodules with the highest ACR TI-RADS point totals that meet criteria for FNA. Size should not be the primary criterion for deciding which nodule(s) to sample. We discourage usage of the term dominant nodule, which is often applied to the largest lesion in the gland, because it downplays the primary role of architecture in determining management. If three or more nodules fall within ACR TI-RADS guidelines for biopsy, the two with the most suspicious appearance on the basis of point totals should be
biopsied, even if they are not the largest. Conversely, in a gland that contains multiple discrete nodules that do not meet criteria, there is little reason to subject the largest one to FNA solely because of its size. Although malignancy cannot be definitively excluded under these conditions, sampling variation lessens the chance of detecting cancer because it would be impractical to biopsy every nodule. Likewise, biopsy is usually not indicated in a gland that is replaced by multiple, confluent nodules of similar appearance.

Assessment of Cervical Lymph Nodes
Evaluation of cervical lymph nodes is a vital part of every thyroid sonographic examination [64]. Abnormal findings suggestive of cervical lymph node metastasis include a globular shape, loss of the normal echogenic hilum, presence of peripheral rather than hilar flow, heterogeneity with cystic components, and punctate echogenic foci that may represent microcalcifications. Detailed discussions of these features and representative ultrasound images are available elsewhere [65-67]. We recommend FNA of suspicious nodes, in addition to up to two nodules that warrant biopsy on the basis of the ACR TI-RADS.

FUTURE DIRECTIONS
The committee believes that the ACR TI-RADS meets our stated goals, although acknowledging the limitations of our additive approach, which does not fully account for the possibility that the risk conferred by a given ultrasound feature may vary depending on what other features are present in a nodule. As well, several committee members have embarked on a parallel project in which interobserver variability of ultrasound feature assignment will be measured. We plan to revise the ACR TI-RADS periodically as results from this study and other evidence comes to light.

TAKE-HOME POINTS
- The goal of this project is to define a risk stratification system for thyroid nodules (ACR TI-RADS) to guide decisions regarding FNA and follow-up.
- The ACR TI-RADS chart allows practitioners to assign points to nodules based on ultrasound features from a standardized lexicon that lends itself to structured reporting.
- The features that form the basis of this system will be assessed for inter-observer variability in a parallel project.

ACKNOWLEDGMENTS
The authors thank Dr Herbert Chen, chair of the Department of Surgery at the University of Alabama at Birmingham, for reviewing the manuscript. We also are grateful to Dr Xuan V. Nguyen of Ohio State University for performing the SEER data analysis.

ADDITIONAL RESOURCES
Additional resources can be found online at: http://dx.doi.org/10.1016/j.jacr.2017.01.046.

REFERENCES


47. Brito JP, Ito Y, Miyauchi A, Tottle RM. A clinical framework to facilitate risk stratification when considering an active surveillance.
alternative to immediate biopsy and surgery in papillary microcarcinoma. Thyroid 2016;26:144-9.


2015 American Thyroid Association Management Guidelines for Adult Patients with Thyroid Nodules and Differentiated Thyroid Cancer

The American Thyroid Association Guidelines Task Force on Thyroid Nodules and Differentiated Thyroid Cancer

Bryan R. Haugen,1,* Erik K. Alexander,2 Keith C. Bible,3 Gerard M. Doherty,4 Susan J. Mandel,5 Yuri E. Nikiforov,6 Furio Pacini,7 Gregory W. Randolph,8 Anna M. Sawka,9 Martin Schlumberger,10 Kathryn G. Schuff,11 Steven I. Sherman,12 Julie Ann Sosa,13 David L. Steward,14 R. Michael Tuttle,15 and Leonard Wartofsky16

Background: Thyroid nodules are a common clinical problem, and differentiated thyroid cancer is becoming increasingly prevalent. Since the American Thyroid Association’s (ATA’s) guidelines for the management of these disorders were revised in 2009, significant scientific advances have occurred in the field. The aim of these guidelines is to inform clinicians, patients, researchers, and health policy makers on published evidence relating to the diagnosis and management of thyroid nodules and differentiated thyroid cancer.

Methods: The specific clinical questions addressed in these guidelines were based on prior versions of the guidelines, stakeholder input, and input of task force members. Task force panel members were educated on knowledge synthesis methods, including electronic database searching, review and selection of relevant citations, and critical appraisal of selected studies. Published English language articles on adults were eligible for inclusion. The American College of Physicians Guideline Grading System was used for critical appraisal of evidence and grading strength of recommendations for therapeutic interventions. We developed a similarly formatted system to appraise the quality of such studies and resultant recommendations. The guideline panel had complete editorial independence from the ATA. Competing interests of guideline task force members were regularly updated, managed, and communicated to the ATA and task force members.

Results: The revised guidelines for the management of thyroid nodules include recommendations regarding initial evaluation, clinical and ultrasound criteria for fine-needle aspiration biopsy, interpretation of fine-needle aspiration biopsy results, use of molecular markers, and management of benign thyroid nodules. Recommendations regarding the initial management of thyroid cancer include those relating to screening for thyroid cancer, staging and risk assessment, surgical management, radioiodine remnant ablation and therapy, and thyrotropin suppression therapy using levothyroxine. Recommendations related to long-term management...
INTRODUCTION

Thyroid nodules are a common clinical problem. Epidemiologic studies have shown the prevalence of palpable thyroid nodules to be approximately 5% in women and 1% in men living in iodine-sufficient parts of the world (1,2). In contrast, high-resolution ultrasound (US) can detect thyroid nodules in 19%–68% of randomly selected individuals, with higher frequencies in women and the elderly (3,4). The clinical importance of thyroid nodules rests with the need to exclude thyroid cancer, which occurs in 7%–15% of cases depending on age, sex, radiation exposure history, family history, and other factors (5,6). Differentiated thyroid cancer (DTC), which includes papillary and follicular cancer, comprises the vast majority (>90%) of all thyroid cancers (7). In the United States, approximately 63,000 new cases of thyroid cancer were predicted to be diagnosed in 2014 (8) compared with 37,200 in 2009 when the last ATA guidelines were published. The yearly incidence has nearly tripled from 4.9 per 100,000 in 1975 to 14.3 per 100,000 in 2009 (9). Almost the entire change has been attributed to an increase in the incidence of papillary thyroid cancer (PTC). Moreover, 25% of the new thyroid cancers diagnosed in 1988–1989 were ≤1 cm compared with 39% of the new thyroid cancer diagnoses in 2008–2009 (9). This tumor shift may be due to the increasing use of neck ultrasonography or other imaging and early diagnosis and treatment (10), trends that are changing the initial treatment and follow-up for many patients with thyroid cancer.

A recent population-based study from Olmsted County reported the doubling of thyroid cancer incidence from 2000 to 2012 compared to the prior decade as entirely attributable to clinically occult cancers detected incidentally on imaging or pathology (11). By 2019, one study predicts that PTC will become the third most common cancer in women at a cost of $19–21 billion in the United States (12). Optimization of long-term health outcomes and education about potential prognosis for individuals with thyroid neoplasms is critically important.

In 1996, the American Thyroid Association (ATA) published treatment guidelines for patients with thyroid nodules and DTC (13). Over the last 15–20 years, there have been many advances in the diagnosis and therapy of both thyroid nodules and DTC, but clinical controversy exists in many areas. A long history of insufficient peer-reviewed research funding for high-quality clinical trials in the field of thyroid neoplasia may be an important contributing factor to existing clinical uncertainties (12). Methodologic limitations or conflicting findings of older studies present a significant challenge to modern-day medical decision-making in many aspects of thyroid neoplasia. Although they are not a specific focus of these guidelines, we recognize that feasibility and cost considerations of various diagnostic and therapeutic options also present important clinical challenges in many clinical practice settings.

AIM AND TARGET AUDIENCE

Our objective in these guidelines is to inform clinicians, patients, researchers, and health policy makers about the best available evidence (and its limitations), relating to the diagnosis and treatment of adult patients with thyroid nodules and DTC. These guidelines should not be applied to children (<18–20 years old); recent ATA guidelines for children with thyroid nodules and DTC were published in 2015 (14). This document is intended to inform clinical decision-making. A major goal of these guidelines is to minimize potential harm from overtreatment in a majority of patients at low risk for disease-specific mortality and morbidity, while appropriately treating and monitoring those patients at higher risk. These guidelines should not be interpreted as a replacement for clinical judgement and should be used to complement informed, shared patient–health care provider deliberation on complex issues. It is important to note that national clinical practice guidelines may not necessarily constitute a legal standard of care in all jurisdictions (15). If important differences in practice settings present barriers to meaningful implementation of the recommendations of these guidelines, interested physicians or groups (in or outside of the United States) may consider adapting the guidelines using established methods (16,17) (ADAPTE Collaboration, 2009; www.g-i-n.net). The ADAPTE Collaboration is an international group of researchers, guideline developers, and guideline implementers who aim to promote the development and use of clinical practice guidelines through the adaption of existing guidelines. Because our primary focus was reviewing the quality of evidence related to health outcomes and diagnostic testing, we decided a priori not to focus on economic resource implications in these guidelines. As part of our review, we identified some knowledge gaps in the field, with associated future research priorities.

Other groups have previously developed clinical practice guidelines, including the American Association of Clinical Endocrinologists, Associazione Medici Endocrinologi, and the European Thyroid Association (18), the British Thyroid Association and The Royal College of Physicians (19), and the National Comprehensive Cancer Network (www.nccn.org). The European Thyroid Association has published consensus guidelines for postoperative US in the management of DTC (20). The Society for Nuclear Medicine and Molecular Imaging (21) and the European Association of Nuclear Medicine have also published guidelines for radioiodine (RAI) therapy of DTC (22). The Japanese Society of Thyroid Surgeons and the Japanese Association of Endocrine Surgeons...
have recently revised guidelines on treatment of patients with thyroid tumors (23). Given the existing controversies in the field, differences in critical appraisal approaches for existing evidence, and differences in clinical practice patterns across geographic regions and physician specialties, it should not be surprising that the organizational guidelines are not in complete agreement for all issues. Such differences highlight the importance of clarifying evidence uncertainties with future high quality clinical research.

**METHODS**

ATA Thyroid Nodules and Differentiated Thyroid Cancer guidelines were published in 2006 (24) and revised in 2009 (25). Because of the rapid growth of the literature on this topic, plans for revising the guidelines within approximately 4 years of publication were made at the inception of the project. A task force chair was appointed by the ATA President with approval of the Board. A task force of specialists with complementary expertise (endocrinology, surgery, nuclear medicine, radiology, pathology, oncology, molecular diagnostics, and epidemiology) was appointed. In order to have broad specialty and geographic representation, as well as fresh perspectives, one-third of the task force is replaced for each iteration of the guidelines, per ATA policy. Upon discussion among the panel members and the Chair with other Chairs of other ATA guideline committees, the American College of Physicians’ (ACP) Grading System was adopted.

**Table 1. Interpretation of the American College of Physicians’ Guideline Grading System (for Therapeutic Interventions)**

<table>
<thead>
<tr>
<th>Recommendation</th>
<th>Clarity of risk/benefit</th>
<th>Implications</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Strong recommendation</strong></td>
<td>Benefits clearly outweigh harms and burdens, or vice versa.</td>
<td><strong>Patients</strong>: Most would want course of action; a person should request discussion if an intervention is not offered. <strong>Clinicians</strong>: Most patients should receive the recommended course of action. <strong>Policymakers</strong>: The recommendation can be adopted as policy in most circumstances.</td>
</tr>
<tr>
<td><strong>Weak recommendation</strong></td>
<td>Benefits closely balanced with harms and burdens.</td>
<td><strong>Patients</strong>: Many would want course of action, but some may not; the decision may depend on individual circumstances. <strong>Clinicians</strong>: Different choices will be appropriate for different patients; the management decision should be consistent with patients’ preferences and circumstances. <strong>Policymakers</strong>: Policymaking will require careful consideration and stakeholder input.</td>
</tr>
<tr>
<td><strong>No recommendation</strong></td>
<td>Balance of benefits and risks cannot be determined.</td>
<td>Decisions based on evidence cannot be made.</td>
</tr>
</tbody>
</table>

**Table 2. Recommendations (for Therapeutic Interventions) Based on Strength of Evidence**

<table>
<thead>
<tr>
<th>Recommendation and evidence quality</th>
<th>Description of supporting evidence</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Strong recommendation</strong></td>
<td>High-quality evidence</td>
<td>RCT without important limitations or overwhelming evidence from observational studies</td>
</tr>
<tr>
<td></td>
<td>Moderate-quality evidence</td>
<td>RCT with important limitations or strong evidence from observational studies</td>
</tr>
<tr>
<td></td>
<td>Low-quality evidence</td>
<td>Observational studies/case studies</td>
</tr>
<tr>
<td><strong>Weak recommendation</strong></td>
<td>High-quality evidence</td>
<td>RCT without important limitations or overwhelming evidence from observational studies</td>
</tr>
<tr>
<td></td>
<td>Moderate-quality evidence</td>
<td>RCT with important limitations or strong evidence from observational studies</td>
</tr>
<tr>
<td></td>
<td>Low-quality evidence</td>
<td>Observational studies/case studies</td>
</tr>
<tr>
<td><strong>Insufficient</strong></td>
<td>Evidence is conflicting, of poor quality, or lacking</td>
<td>Insufficient evidence to recommend for or against</td>
</tr>
</tbody>
</table>

*This description of supporting evidence refers to therapy, therapeutic strategy, or prevention studies. The description of supporting evidence is different for diagnostic accuracy studies. RCT, randomized controlled trial.*
for use in these guidelines, relating to critical appraisal and recommendations on therapeutic interventions (26) (Tables 1 and 2). An important component of these guidelines was judged to be critical appraisal of studies of diagnostic tests; however, the ACP Guideline Grading System is not designed for this purpose. We reviewed a number of appraisal systems for diagnostic tests, but some of the complexity and the time-consuming nature of some systems limited their feasibility for implementation in our group (27–31). We drafted, revised, and piloted the use of a newly developed diagnostic test appraisal system that was acceptable to panel members. This system included consideration of the following

<table>
<thead>
<tr>
<th>Recommendation</th>
<th>Accuracy of diagnostic information versus risks and burden of testing</th>
<th>Implications</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Strong recommendation</strong></td>
<td>Knowledge of the diagnostic test result clearly outweighs risks and burden of testing or vice versa.</td>
<td><strong>Patients:</strong> In the case of an accurate test for which benefits outweigh risks/burden, most would want the diagnostic to be offered (with appropriate counseling). A patient should request discussion of the test if it is not offered. In contrast, for a test in which risks and burden outweigh the benefits, most patients should not expect the test to be offered. <strong>Clinicians:</strong> In the case of an accurate test for which benefits outweigh risks/burden, most patients should be offered the diagnostic test (and provided relevant counseling). Counseling about the test should include a discussion of the risks, benefits, and uncertainties related to testing (as applicable), as well as the implications of the test result. In contrast, for a test in which risks and burden outweigh the perceived benefits, most patients should not be offered the test, or if the test is discussed, the rationale against the test should, for the particular clinical situation, be explained. <strong>Policymakers:</strong> In the case of an accurate test for which benefits outweigh risks/burden, availability of the diagnostic test should be adopted in health policy. In contrast, for a test in which risks and burden outweigh the perceived benefits, some restrictions on circumstances for test use may need to be considered.</td>
</tr>
<tr>
<td><strong>Weak recommendation</strong></td>
<td>Knowledge of the diagnostic test result is closely balanced with risks and burden of testing.</td>
<td><strong>Patients:</strong> Most would want to be informed about the diagnostic test, but some would not want to seriously consider undergoing the test; a decision may depend on the individual circumstances (e.g., risk of disease, comorbidities, or other), the practice environment, feasibility of optimal execution of the test, and consideration of other available options. <strong>Clinicians:</strong> Different choices will be appropriate for different patients, and counseling about the test (if being considered) should include a discussion of the risks, benefits, and uncertainties related to testing (as applicable), as well as the implications of the test result. The decision to perform the test should include consideration of the patients’ values, preferences, feasibility, and the specific circumstances. Counseling the patient on why the test may be helpful or not, in her/his specific circumstance, may be very valuable in the decision-making process. <strong>Policymakers:</strong> Policymaking decisions on the availability of the test will require discussion and stakeholder involvement.</td>
</tr>
<tr>
<td><strong>No recommendation</strong></td>
<td>Balance of knowledge of the diagnostic test result cannot be determined.</td>
<td>Decisions on the use of the test based on evidence from scientific studies cannot be made.</td>
</tr>
</tbody>
</table>

*Frequently in these guidelines, the accuracy of the diagnosis of thyroid cancer (relative to a histologic gold standard) was the diagnostic outcome unless otherwise specified. However, prognostic, disease staging, or risk stratification studies were also included in the grading scheme of diagnostic studies. For disease staging systems, the implication for use would be on the part of the clinician, in reporting results in the medical record and communicating them to the patient (at the applicable time point in disease or follow-up trajectory), as opposed to offering a specific choice of staging/risk stratification system to the patient.*
methodologic elements: consecutive recruitment of patients representative of clinical practice, use of an appropriate reference gold standard, directness of evidence (e.g., target population of interest, testing procedures representative of clinical practice, and relevant outcomes), precision of diagnostic accuracy measures (e.g., width of confidence intervals for estimates such as sensitivity, specificity), and consistency of results among studies using the same test (Tables 3 and 4). In the majority of circumstances (unless otherwise specified), the outcome of interest for the diagnostic test was the diagnosis of thyroid cancer (relative to a histologic gold standard). However, prognostic studies were also graded using the diagnostic study critical appraisal framework. In terms of strength of recommendation for use of diagnostic studies, we modeled our approach on the ACP system for therapeutic studies, as previously described, but the target outcome was the accuracy in establishing a definitive diagnosis, largely relating to the diagnosis of new or recurrent malignancy (unless otherwise specified). Diagnostic tests or risk stratification systems used for estimation of prognosis were also appraised using the diagnostic test grading system. An important limitation of our diagnostic test appraisal system is that it does not specifically examine the clinical utility of a test in improving long-term health outcomes by execution of the test as part of an intended therapeutic strategy (unless specifically noted). However, as much as possible, we tried to separate recommendations on the diagnostic accuracy of a test from therapeutic management based on the test result, with the latter grading being more rigorous and based on longer term outcomes (whenever possible). It is important to note that according to our diagnostic test grading system, a body of well-executed nonrandomized diagnostic accuracy

<table>
<thead>
<tr>
<th>Recommendation and evidence quality</th>
<th>Methodologic quality of supporting evidence</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strong recommendation</td>
<td>Evidence from one or more well-designed nonrandomized diagnostic accuracy studies (i.e., observational—cross-sectional or cohort) or systematic reviews/meta-analyses of such observational studies (with no concern about internal validity or external generalizability of the results)</td>
<td>Implies the test can be offered to most patients in most applicable circumstances without reservation.</td>
</tr>
<tr>
<td>Moderate-quality evidence</td>
<td>Evidence from nonrandomized diagnostic accuracy studies (cross-sectional or cohort), with one or more possible limitations causing minor concern about internal validity or external generalizability of the results</td>
<td>Implies the test can be offered to most patients in most applicable circumstances without reservation.</td>
</tr>
<tr>
<td>Low-quality evidence</td>
<td>Evidence from nonrandomized diagnostic accuracy studies with one or more important limitations causing serious concern about internal validity or external generalizability of the results</td>
<td>Implies the test can be offered to most patients in most applicable circumstances, but the utilization of the test may change when higher-quality evidence becomes available.</td>
</tr>
<tr>
<td>Weak recommendation</td>
<td>Evidence from one or more well-designed nonrandomized diagnostic accuracy studies (i.e., observational—cross-sectional or cohort) or systematic reviews/meta-analyses of such observational studies (with no concern about internal validity or external generalizability of the results)</td>
<td>The degree to which the diagnostic test is seriously considered may differ depending on circumstances or patients’ or societal values.</td>
</tr>
<tr>
<td>Moderate-quality evidence</td>
<td>Evidence from nonrandomized diagnostic accuracy studies (cross-sectional or cohort), with one or more possible limitations causing minor concern about internal validity or external generalizability of the results</td>
<td>The degree to which the diagnostic test is seriously considered may differ depending on individual patients' practice circumstances or patients’ or societal values.</td>
</tr>
<tr>
<td>Low-quality evidence</td>
<td>Evidence from nonrandomized diagnostic accuracy studies with one or more important limitations causing serious concern about internal validity or external generalizability of the results.</td>
<td>Alternative options may be equally reasonable.</td>
</tr>
<tr>
<td>Insufficient</td>
<td>Evidence may be of such poor quality, conflicting, lacking (i.e., studies not done), or not externally generalizable to the target clinical population such that the estimate of the true effect of the test is uncertain and does not permit a reasonable conclusion to be made.</td>
<td>Insufficient evidence exists to recommend for or against routinely offering the diagnostic test.</td>
</tr>
</tbody>
</table>
studies could be considered high-quality evidence; yet, a therapeutic strategy incorporating the use of the diagnostic test would require one or more well-executed randomized controlled trials (RCTs) to be considered high-quality evidence. In developing and applying our diagnostic test critical appraisal system, we considered American societal values, relating to the importance of informing patients about potentially helpful tests developed for their clinical situation (with counseling on relevant limitations) and the role of patients in informed, shared decision-making relating to diagnostic and therapeutic strategies. Such input was based on thoughtful consideration of stakeholder input, including input from physician stakeholders who were committee members. Because this was a preliminary pilot utilization of this diagnostic test critical appraisal system by our group, we have labeled recommendations using this system in the manuscript (diagnostic test recommendation). Moreover, we anticipate that the future iterations of these guidelines will likely incorporate further refinements to the system, or even possible adoption of another system, if it is superior and feasible to execute by contributing physicians.

Prior to initiating the reviews, all task force members were provided written and verbal group advice on conducting electronic literature searches, critical appraisal of articles, and rationale for formulating strength of recommendations from a panel member with epidemiology and systematic review expertise (via e-mail documents, a teleconference meeting on February 21, 2012). For each question, a primary reviewer performed a literature search, appraised relevant literature, generated recommendations, accompanying text, and a relevant bibliography. This was then reviewed by the secondary reviewer, revised as needed, and presented for review by the entire panel. Feedback and suggestions for revisions from the Chair and panel members were obtained via e-mail, regularly scheduled teleconferences, and face-to-face meetings held in conjunction with scientific meetings. Once the manuscript was drafted, all suggestions for revisions were regularly reviewed by all panel members in the form of a tracked changes draft manuscript and teleconferences. The draft document continued to be revised until no further suggestions for further revisions were requested by any panel members. Thus, general consensus on acceptability of recommendations and manuscript text was achieved, with the fundamental understanding that not all recommendations may be feasible in all practice settings.

Formal stakeholder input in development of these guidelines was sought from ATA membership in an online survey distributed in October 2011. Thyroid cancer survivor group leadership input was sought from three North American thyroid cancer groups via e-mail correspondence in January to March of 2012. We also reviewed any letters, editorials, or reviews of the 2009 iteration of the guidelines (25) that were collected by the current Chair of the committee. Pre-publication verbal feedback on some of the key guideline recommendations was received at a formal Satellite Symposium held in conjunction with the Endocrine Society meeting in Chicago on June 19, 2014. The guideline manuscript was reviewed and approved by the ATA Board of Directors, then made available to the ATA membership for review and comments in September 2014. Substantive comments were received from 33 members representing endocrinology, surgery, pathology, and nuclear medicine. Feedback and suggestions were formally discussed by the panel, and revisions were made to the manuscript prior to journal submission. The organization of management guidelines recommendations is shown in Table 5.

The medical opinions expressed here are those of the authors, and the committee had complete editorial independence from the ATA in writing the guidelines. No funding was received by individual committee members from the ATA or industry for work on these guidelines. Competing interests of all committee members were reviewed at inception of the group, yearly, and upon completion of the guidelines and are included with this document.

Table 5. Organization of the 2015 ATA Guidelines for Thyroid Nodules and Differentiated Thyroid Cancer

<table>
<thead>
<tr>
<th>Page</th>
<th>Location key</th>
<th>Sections and subsections</th>
<th>Item</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>[A1]</td>
<td><strong>THYROID NODULE GUIDELINES</strong></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>[A2]</td>
<td>What is the role of thyroid cancer screening in people with familial follicular cell–derived DTC?</td>
<td>R1</td>
</tr>
<tr>
<td>10</td>
<td>[A3]</td>
<td>What is the appropriate laboratory and imaging evaluation for patients with clinically or incidentally discovered thyroid nodules?</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>[A4]</td>
<td>Serum thyrotropin measurement</td>
<td>R2</td>
</tr>
<tr>
<td>11</td>
<td>[A5]</td>
<td>Serum thyroglobulin measurement</td>
<td>R3</td>
</tr>
<tr>
<td>11</td>
<td>[A6]</td>
<td>Serum calcitonin measurement</td>
<td>R4</td>
</tr>
<tr>
<td>11</td>
<td>[A7]</td>
<td>[18F]Fluorodeoxyglucose positron emission tomography</td>
<td>R5</td>
</tr>
<tr>
<td>12</td>
<td>[A8]</td>
<td>Thyroid sonography</td>
<td>R6</td>
</tr>
<tr>
<td>12</td>
<td>[A9]</td>
<td>US for FNA decision-making</td>
<td>R7</td>
</tr>
<tr>
<td>12</td>
<td>[A10]</td>
<td>Recommendations for diagnostic FNA of a thyroid nodule based on sonographic pattern</td>
<td>R8c, F1c, F2c, T6c</td>
</tr>
<tr>
<td>16</td>
<td>[A11]</td>
<td>What is the role of FNA, cytology interpretation, and molecular testing in patients with thyroid nodules?</td>
<td>R9c, F1c, T7c</td>
</tr>
<tr>
<td>17</td>
<td>[A12]</td>
<td>Nondiagnostic cytology</td>
<td>R10</td>
</tr>
<tr>
<td>17</td>
<td>[A13]</td>
<td>Benign cytology</td>
<td>R11</td>
</tr>
<tr>
<td>18</td>
<td>[A14]</td>
<td>Malignant cytology</td>
<td>R12</td>
</tr>
<tr>
<td>19</td>
<td>[A15]</td>
<td>Indeterminate cytology (AUS/FLUS, FN, SUSP)</td>
<td></td>
</tr>
</tbody>
</table>

(continued)
<table>
<thead>
<tr>
<th>Page</th>
<th>Location key</th>
<th>Sections and subsections</th>
<th>Item</th>
</tr>
</thead>
<tbody>
<tr>
<td>19</td>
<td>[A16]</td>
<td><strong>What are the principles of the molecular testing of FNA samples?</strong>&lt;sup&gt;b&lt;/sup&gt;</td>
<td>R13–14</td>
</tr>
<tr>
<td>21</td>
<td>[A17]</td>
<td>AUS/FLUS cytology&lt;sup&gt;c&lt;/sup&gt;</td>
<td>R15&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>22</td>
<td>[A18]</td>
<td>Follicular neoplasm/suspicious for follicular neoplasm cytology&lt;sup&gt;c&lt;/sup&gt;</td>
<td>R16&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>23</td>
<td>[A19]</td>
<td>Suspicious for malignancy cytology&lt;sup&gt;c&lt;/sup&gt;</td>
<td>R17&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>23</td>
<td>[A20]</td>
<td><strong>What is the utility of</strong>&lt;sup&gt;a&lt;/sup&gt; FDG-PET scanning to predict malignant or benign disease when FNA cytology is indeterminate (AUS/FLUS, FN, SUSP)?&lt;sup&gt;b&lt;/sup&gt;</td>
<td>R18&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>23</td>
<td>[A21]</td>
<td><strong>What is the appropriate operation for cytologically indeterminate thyroid nodules?</strong>&lt;sup&gt;c&lt;/sup&gt;</td>
<td>R19–20&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>25</td>
<td>[A22]</td>
<td><strong>How should multinodular thyroid glands (i.e., two or more clinically relevant nodules) be evaluated for malignancy?</strong></td>
<td>R21–22</td>
</tr>
<tr>
<td>25</td>
<td>[A23]</td>
<td><strong>What are the best methods for long-term follow-up of patients with thyroid nodules?</strong></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>[A24]</td>
<td>Recommendations for initial follow-up of nodules with benign FNA cytology&lt;sup&gt;c&lt;/sup&gt;</td>
<td>R23A–C&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>25</td>
<td>[A25]</td>
<td>Recommendation for follow-up of nodules with two benign FNA cytology results&lt;sup&gt;b&lt;/sup&gt;</td>
<td>R23D&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>26</td>
<td>[A26]</td>
<td>Follow-up for nodules that do not meet FNA criteria&lt;sup&gt;b&lt;/sup&gt;</td>
<td>R24&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>27</td>
<td>[A27]</td>
<td><strong>What is the role of medical or surgical therapy for benign thyroid nodules?</strong></td>
<td>R25–29</td>
</tr>
<tr>
<td>27</td>
<td>[A28]</td>
<td><strong>How should thyroid nodules in pregnant women be managed?</strong></td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>[A29]</td>
<td>FNA for thyroid nodules discovered during pregnancy</td>
<td>R30</td>
</tr>
<tr>
<td>28</td>
<td>[A30]</td>
<td>Approaches to pregnant patients with malignant or indeterminate cytology</td>
<td>R31</td>
</tr>
<tr>
<td>28</td>
<td>[B1]</td>
<td><strong>DIFFERENTIATED THYROID CANCER: INITIAL MANAGEMENT GUIDELINES</strong></td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>[B2]</td>
<td>Goals of initial therapy of DTC</td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>[B3]</td>
<td><strong>What is the role of preoperative staging with diagnostic imaging and laboratory tests?</strong></td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>[B4]</td>
<td>Neck imaging—ultrasound</td>
<td>R32 F3, T6, T8&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>30</td>
<td>[B5]</td>
<td>Neck imaging—CT/MRI/PET&lt;sup&gt;c&lt;/sup&gt;</td>
<td>R33&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>31</td>
<td>[B6]</td>
<td>Measurement of serum Tg and anti-Tg antibodies</td>
<td>R34</td>
</tr>
<tr>
<td>31</td>
<td>[B7]</td>
<td>Operative approach for a biopsy diagnostic for follicular cell-derived malignancy&lt;sup&gt;c&lt;/sup&gt;</td>
<td>R35&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>33</td>
<td>[B8]</td>
<td>Lymph node dissection</td>
<td>R36–37, F3</td>
</tr>
<tr>
<td>35</td>
<td>[B9]</td>
<td>Completion thyroidectomy</td>
<td>R38</td>
</tr>
<tr>
<td>35</td>
<td>[B10]</td>
<td><strong>What is the appropriate perioperative approach to voice and parathyroid issues?</strong>&lt;sup&gt;b&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>35</td>
<td>[B11]</td>
<td>Preoperative care communication&lt;sup&gt;b&lt;/sup&gt;</td>
<td>R39&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>35</td>
<td>[B12]</td>
<td>Preoperative voice assessment&lt;sup&gt;c&lt;/sup&gt;</td>
<td>R40–41&lt;sup&gt;b&lt;/sup&gt;, T9&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>36</td>
<td>[B13]</td>
<td>Intraoperative voice&lt;sup&gt;c&lt;/sup&gt; and parathyroid management&lt;sup&gt;b&lt;/sup&gt;</td>
<td>R42–43&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>37</td>
<td>[B14]</td>
<td>Postoperative care&lt;sup&gt;b&lt;/sup&gt;</td>
<td>R44&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>37</td>
<td>[B15]</td>
<td><strong>What are the basic principles of histopathologic evaluation of thyroidectomy samples?</strong>&lt;sup&gt;b&lt;/sup&gt;</td>
<td>R46&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>40</td>
<td>[B16]</td>
<td><strong>What is the role of postoperative staging systems and risk stratification in the management of DTC?</strong></td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>[B17]</td>
<td>Postoperative staging</td>
<td>R47</td>
</tr>
<tr>
<td>40</td>
<td>[B18]</td>
<td>AJCC/UICC TNM staging</td>
<td>T10</td>
</tr>
<tr>
<td>41</td>
<td>[B19]</td>
<td><strong>What initial stratification system should be used to estimate the risk of persistent/recurrent disease?</strong>&lt;sup&gt;c&lt;/sup&gt;</td>
<td>R48&lt;sup&gt;c&lt;/sup&gt;, T11&lt;sup&gt;b&lt;/sup&gt;, T12&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>43</td>
<td>[B20]</td>
<td>Potential impact of specific clinicopathologic features on the risk estimates in PTC&lt;sup&gt;c&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>44</td>
<td>[B21]</td>
<td>Potential impact of <strong>BRAF</strong>&lt;sup&gt;V600E&lt;/sup&gt; and other mutations on risk of estimates in PTC&lt;sup&gt;c&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>45</td>
<td>[B22]</td>
<td>Potential impact of postoperative serum Tg on risk estimates&lt;sup&gt;b&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>46</td>
<td>[B23]</td>
<td>Proposed modifications to the 2009 ATA initial risk stratification system&lt;sup&gt;b&lt;/sup&gt;</td>
<td>T12&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>46</td>
<td>[B24]</td>
<td>Risk of recurrence as a continuum of risk&lt;sup&gt;b&lt;/sup&gt;</td>
<td>F4&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>46</td>
<td>[B25]</td>
<td><strong>How should initial risk estimates be modified over time?</strong>&lt;sup&gt;b&lt;/sup&gt;</td>
<td>R49</td>
</tr>
<tr>
<td>47</td>
<td>[B26]</td>
<td>Proposed terminology to classify response to therapy and clinical</td>
<td></td>
</tr>
</tbody>
</table>

(continued)
<table>
<thead>
<tr>
<th>Page</th>
<th>Location key</th>
<th>Sections and subsections</th>
<th>Item&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>47</td>
<td>[B27]</td>
<td>Implications&lt;sup&gt;b&lt;/sup&gt; Integrated treatment outcomes: no structural, biochemical, or clinical evidence of disease after initial therapy (remission, NED)&lt;sup&gt;b&lt;/sup&gt;</td>
<td>T13&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>50</td>
<td>[B28]</td>
<td>Biochemical incomplete response: abnormal Tg values in the absence of localizable disease&lt;sup&gt;b&lt;/sup&gt;</td>
<td>T13&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>51</td>
<td>[B29]</td>
<td>Structural incomplete response: persistent or newly identified loco-regional or distant metastases&lt;sup&gt;b&lt;/sup&gt;</td>
<td>T13&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>52</td>
<td>[B30]</td>
<td>Indeterminate response: biochemical or structural findings that cannot be classified as either benign or malignant (acceptable response)&lt;sup&gt;b&lt;/sup&gt;</td>
<td>T13&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>52</td>
<td>[B31]</td>
<td>Using risk stratification to guide disease surveillance and therapeutic management decisions&lt;sup&gt;b&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>53</td>
<td>[B32]</td>
<td>Should postoperative disease status be considered in decision-making for RAI therapy for patients with DTC?</td>
<td>R50</td>
</tr>
<tr>
<td>53</td>
<td>[B33]</td>
<td>Utility of postoperative serum Tg in clinical decision-making</td>
<td></td>
</tr>
<tr>
<td>54</td>
<td>[B34]</td>
<td>Potential role of postoperative US in conjunction with postoperative serum Tg in clinical decision-making</td>
<td></td>
</tr>
<tr>
<td>54</td>
<td>[B35]</td>
<td>Role of postoperative radioisotope diagnostic scanning in clinical decision-making</td>
<td></td>
</tr>
<tr>
<td>55</td>
<td>[B36]</td>
<td>What is the role of RAI (including remnant ablation, adjuvant therapy, or therapy persistent disease) after thyroidectomy in the primary management of differentiated thyroid cancer?</td>
<td>R51–56&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>58</td>
<td>[B37]</td>
<td>What is the role of molecular marker status in therapeutic RAI decision-making?</td>
<td>R52&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>58</td>
<td>[B38]</td>
<td>How long does thyroid hormone need to be withdrawn in preparation for RAI remnant ablation/treatment or diagnostic scanning?</td>
<td>R53</td>
</tr>
<tr>
<td>59</td>
<td>[B39]</td>
<td>Can rhTSH (Thyrogen) be used as an alternative to thyroxine withdrawal for remnant ablation or adjuvant therapy in patients who have undergone near-total or total thyroidectomy?</td>
<td>R54</td>
</tr>
<tr>
<td>60</td>
<td>[B40]</td>
<td>What activity of &lt;sup&gt;131&lt;/sup&gt;I should be used for remnant ablation or adjuvant therapy?</td>
<td>R55–56&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>63</td>
<td>[B41]</td>
<td>Is a low-iodine diet necessary before remnant ablation?</td>
<td>R57</td>
</tr>
<tr>
<td>63</td>
<td>[B42]</td>
<td>Should a posttherapy scan be performed following remnant ablation or adjuvant therapy?</td>
<td>R58</td>
</tr>
<tr>
<td>64</td>
<td>[B43]</td>
<td>Early management of DTC after initial therapy</td>
<td></td>
</tr>
<tr>
<td>64</td>
<td>[B44]</td>
<td>What is the appropriate degree of initial TSH suppression?</td>
<td>R59</td>
</tr>
<tr>
<td>65</td>
<td>[B45]</td>
<td>Is there a role for adjunctive external beam radiation or chemotherapy?</td>
<td></td>
</tr>
<tr>
<td>65</td>
<td>[B46]</td>
<td>External beam radiation</td>
<td>R60</td>
</tr>
<tr>
<td>65</td>
<td>[B47]</td>
<td>Systemic adjuvant therapy</td>
<td>R61</td>
</tr>
<tr>
<td>65</td>
<td>[C1]</td>
<td>DTC: LONG-TERM MANAGEMENT AND ADVANCED CANCER MANAGEMENT GUIDELINES</td>
<td></td>
</tr>
<tr>
<td>65</td>
<td>[C2]</td>
<td>What are the appropriate features of long-term management?</td>
<td></td>
</tr>
<tr>
<td>66</td>
<td>[C3]</td>
<td>What are the criteria for absence of persistent tumor (excellent response)?</td>
<td></td>
</tr>
<tr>
<td>66</td>
<td>[C4]</td>
<td>What are the appropriate methods for following patients after initial therapy?</td>
<td></td>
</tr>
<tr>
<td>66</td>
<td>[C5]</td>
<td>What is the role of serum Tg measurement in the follow-up of DTC?</td>
<td>R62–63&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>66</td>
<td>[C6]</td>
<td>Serum Tg measurement and clinical utility</td>
<td></td>
</tr>
<tr>
<td>67</td>
<td>[C7]</td>
<td>Anti-Tg antibodies</td>
<td></td>
</tr>
<tr>
<td>68</td>
<td>[C8]</td>
<td>What is the role of serum Tg measurement in patients who have not undergone RAI remnant ablation?</td>
<td>R64</td>
</tr>
<tr>
<td>69</td>
<td>[C9]</td>
<td>What is the role of US and other imaging techniques (RAI SPECT/CT, CT, MRI, PET-CT) during follow-up?</td>
<td></td>
</tr>
<tr>
<td>69</td>
<td>[C10]</td>
<td>Cervical ultrasonography</td>
<td>R65</td>
</tr>
<tr>
<td>69</td>
<td>[C11]</td>
<td>Diagnostic whole-body RAI scans</td>
<td>R66–67</td>
</tr>
<tr>
<td>70</td>
<td>[C12]</td>
<td>&lt;sup&gt;18&lt;/sup&gt;FDG-PET scanning</td>
<td>R68</td>
</tr>
<tr>
<td>71</td>
<td>[C13]</td>
<td>CT and MRI&lt;sup&gt;b&lt;/sup&gt;</td>
<td>R69&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>72</td>
<td>[C14]</td>
<td>Using ongoing risk stratification (response to therapy) to guide disease long-term surveillance and therapeutic management decisions&lt;sup&gt;c&lt;/sup&gt;</td>
<td></td>
</tr>
</tbody>
</table>

(continued)
### Table 5. (Continued)

<table>
<thead>
<tr>
<th>Page</th>
<th>Location key</th>
<th>Sections and subsections</th>
<th>Item(^a)</th>
</tr>
</thead>
<tbody>
<tr>
<td>72</td>
<td>[C15]</td>
<td>What is the role of TSH suppression during thyroid hormone therapy in the long-term follow-up of DTC?(^c)</td>
<td>R70(^c)</td>
</tr>
<tr>
<td>74</td>
<td>[C16]</td>
<td>What is the most appropriate management of DTC patients with metastatic disease?</td>
<td>T15(^b)</td>
</tr>
<tr>
<td>74</td>
<td>[C17]</td>
<td>What is the optimal directed approach to patients with suspected structural neck recurrence?</td>
<td>R71</td>
</tr>
<tr>
<td>74</td>
<td>[C18]</td>
<td>Nodal size threshold</td>
<td></td>
</tr>
<tr>
<td>75</td>
<td>[C19]</td>
<td>Extent of nodal surgery</td>
<td></td>
</tr>
<tr>
<td>75</td>
<td>[C20]</td>
<td>Ethanol injection(^b)</td>
<td></td>
</tr>
<tr>
<td>75</td>
<td>[C21]</td>
<td>Radiofrequency or laser ablation(^b)</td>
<td></td>
</tr>
<tr>
<td>75</td>
<td>[C22]</td>
<td>Other therapeutic options(^b)</td>
<td></td>
</tr>
<tr>
<td>76</td>
<td>[C23]</td>
<td>What is the surgical management of aerodigestive invasion?</td>
<td>R72</td>
</tr>
<tr>
<td>76</td>
<td>[C24]</td>
<td>How should RAI therapy be considered for loco-regional or distant metastatic disease?</td>
<td></td>
</tr>
<tr>
<td>76</td>
<td>[C25]</td>
<td>Administered activity of 131I for loco-regional or metastatic disease(^c)</td>
<td>R73(^c)</td>
</tr>
<tr>
<td>77</td>
<td>[C26]</td>
<td>Use of rhTSH (Thyrogen) to prepare patients for 131I therapy for loco-regional or metastatic disease</td>
<td>R74–75</td>
</tr>
<tr>
<td>77</td>
<td>[C27]</td>
<td>Use of lithium in 131I therapy</td>
<td>R76</td>
</tr>
<tr>
<td>77</td>
<td>[C28]</td>
<td>How should distant metastatic disease to various organs be treated?</td>
<td></td>
</tr>
<tr>
<td>78</td>
<td>[C29]</td>
<td>Treatment of pulmonary metastases</td>
<td>R77–78</td>
</tr>
<tr>
<td>78</td>
<td>[C30]</td>
<td>RAI treatment of bone metastases</td>
<td>R79</td>
</tr>
<tr>
<td>79</td>
<td>[C31]</td>
<td>When should empiric RAI therapy be considered for Tg-positive, RAI diagnostic scan–negative patients?</td>
<td>R80–82</td>
</tr>
<tr>
<td>79</td>
<td>[C32]</td>
<td>What is the management of complications of RAI therapy?</td>
<td>R83–85</td>
</tr>
<tr>
<td>80</td>
<td>[C33]</td>
<td>How should patients who have received RAI therapy be monitored for risk of secondary malignancies?</td>
<td>R86</td>
</tr>
<tr>
<td>80</td>
<td>[C34]</td>
<td>What other testing should patients receiving RAI therapy undergo?</td>
<td>R87</td>
</tr>
<tr>
<td>80</td>
<td>[C35]</td>
<td>How should patients be counseled about RAI therapy and pregnancy, breastfeeding, and gonadal function?</td>
<td>R88–90</td>
</tr>
<tr>
<td>81</td>
<td>[C36]</td>
<td>How is RAI-refractory DTC classified?(^b)</td>
<td>R91(^b)</td>
</tr>
<tr>
<td>82</td>
<td>[C37]</td>
<td>Which patients with metastatic thyroid cancer can be followed without additional therapy?(^b)</td>
<td>R92(^b)</td>
</tr>
<tr>
<td>82</td>
<td>[C38]</td>
<td>What is the role for directed therapy in advanced thyroid cancer?(^c)</td>
<td>R93(^c)</td>
</tr>
<tr>
<td>84</td>
<td>[C39]</td>
<td>Treatment of brain metastases</td>
<td>R94</td>
</tr>
<tr>
<td>84</td>
<td>[C40]</td>
<td>Who should be considered for clinical trials?(^b)</td>
<td>R95(^b)</td>
</tr>
<tr>
<td>84</td>
<td>[C41]</td>
<td>What is the role of systemic therapy (kinase inhibitors, other selective therapies, conventional chemotherapy, bisphosphonates) in treating metastatic DTC?(^b)</td>
<td></td>
</tr>
<tr>
<td>85</td>
<td>[C42]</td>
<td>Kinase inhibitors(^b)</td>
<td>R96(^b), T16(^b)</td>
</tr>
<tr>
<td>87</td>
<td>[C43]</td>
<td>Patients for whom first-line kinase inhibitor therapy fails(^b)</td>
<td>R97(^b), T17(^b)</td>
</tr>
<tr>
<td>87</td>
<td>[C44]</td>
<td>Management of toxicities from kinase inhibitor therapy(^b)</td>
<td>R98(^b), T17(^b)</td>
</tr>
<tr>
<td>87</td>
<td>[C45]</td>
<td>Other novel agents(^b)</td>
<td>R99</td>
</tr>
<tr>
<td>87</td>
<td>[C46]</td>
<td>Cytotoxic chemotherapy</td>
<td>R100</td>
</tr>
<tr>
<td>88</td>
<td>[C47]</td>
<td>Bone-directed agents(^c)</td>
<td>R101(^c)</td>
</tr>
<tr>
<td>89</td>
<td>[D1]</td>
<td>DIRECTIONS FOR FUTURE RESEARCH</td>
<td></td>
</tr>
<tr>
<td>89</td>
<td>[D2]</td>
<td>Optimizing molecular markers for diagnosis, prognosis, and therapeutic targets</td>
<td></td>
</tr>
<tr>
<td>89</td>
<td>[D3]</td>
<td>Active surveillance of DTC primary tumors</td>
<td></td>
</tr>
<tr>
<td>89</td>
<td>[D4]</td>
<td>Improved risk stratification</td>
<td></td>
</tr>
<tr>
<td>90</td>
<td>[D5]</td>
<td>Improving our understanding of the risks and benefits of DTC treatments and optimal implementation/utilization</td>
<td></td>
</tr>
<tr>
<td>90</td>
<td>[D6]</td>
<td>Issues with measurement of Tg and anti-Tg antibodies</td>
<td></td>
</tr>
<tr>
<td>90</td>
<td>[D7]</td>
<td>Management of metastatic cervical adenopathy detected on US</td>
<td></td>
</tr>
<tr>
<td>91</td>
<td>[D8]</td>
<td>Novel therapies for systemic RAI-refractory disease</td>
<td></td>
</tr>
<tr>
<td>91</td>
<td>[D9]</td>
<td>Survivorship care</td>
<td></td>
</tr>
</tbody>
</table>

\(^a\)F, figure; R, recommendation; T, table.

\(^b\)New section/recommendation.

\(^c\)Substantially changed recommendation compared with 2009.

ATA, American Thyroid Association; AUS/FLUS, atypia of undetermined significance/follicular lesion of undetermined significance; CT, computed tomography; DTC, differentiated thyroid cancer; FN, follicular neoplasm; FNA, fine-needle aspiration; \(^{18}\)FDG-PET, \(^{18}\)fluorodeoxyglucose positron emission tomography; MRI, magnetic resonance imaging; NED, no evidence of disease; PET, positron emission tomography; RAI, radioactive iodine (radioiodine); rhTSH, recombinant human thyrotropin; SPECT/CT, single photon emission computed tomography–computed tomography; SUSP, suspicious for malignancy; Tg, thyroglobulin; TSH, thyrotropin; US, ultrasound.
[A1] THYROID NODULE GUIDELINES

A thyroid nodule is a discrete lesion within the thyroid gland that is radiologically distinct from the surrounding thyroid parenchyma. Some palpable lesions may not correspond to distinct radiologic abnormalities (32). Such abnormalities do not meet the strict definition for thyroid nodules. Nonpalpable nodules detected on US or other anatomic imaging studies are termed incidentally discovered nodules or “incidentalomas.” Nonpalpable nodules have the same risk of malignancy as do sonographically confirmed palpable nodules of the same size (33). Generally, only nodules >1 cm should be evaluated, since they have a greater potential to be clinically significant cancers. Occasionally, there may be nodules <1 cm that require further evaluation because of clinical symptoms or associated lymphadenopathy. In very rare cases, some nodules <1 cm lack these sonographic and clinical warning signs yet may nonetheless cause future morbidity and mortality. This remains highly unlikely, and given the unfavorable cost/benefit considerations, attempts to diagnose and treat all such small thyroid cancers in an effort to prevent exceedingly rare outcomes is deemed to cause more harm than good. In general, the guiding clinical strategy acknowledges that most thyroid nodules are low risk, and many thyroid cancers pose minimal risk to human health and can be effectively treated.

[A2] What is the role of thyroid cancer screening in people with familial follicular cell–derived DTC?

- **RECOMMENDATION 1**

  Screening people with familial follicular cell–derived DTC may lead to an earlier diagnosis of thyroid cancer, but the panel cannot recommend for or against US screening since there is no evidence that this would lead to reduced morbidity or mortality.

  (No recommendation, Insufficient evidence)

  Screening programs for patients at risk of oncological disease are usually advocated based on the following evidence: (a) a clear demonstration that the patient is indeed at risk; (b) demonstration that screening allows the detection of the disease at an earlier stage; (c) early diagnosis has an impact on subsequent outcome, both recurrence and survival.

  Family members of patients with nonmedullary DTC may be considered at risk based on epidemiological evidence showing that 5%–10% of DTCs have a familial occurrence. However, in most of the pedigrees only two members are affected. There is controversy on whether two family members are sufficient to define a real familial disease rather than a fortuitous association. The probability estimates reported by Charkes (34) suggests that when only two first-degree family members are affected, the probability that the disease is sporadic is 62%. This probability decreases when the number of affected family members is three or more. In contrast, the study by Capezzone et al. (35), which was statistically adjusted to minimize risk of “insufficient follow-up bias,” demonstrates that even when only two family members are affected, the disease displays the features of “genetic anticipation” (occurrence of the disease at an earlier age and with more aggressive presentation in the subsequent generation compared with the first generation), which is considered good evidence for a distinct clinical entity possibly representing true familial disease. Appearance of the disease at an earlier age has also been found by Moses et al. (36). More advanced disease at presentation and slightly worse outcomes have been reported in familial cases by Capezzone et al. (35). More frequent multicentricity has been reported by Ito et al. (37), but disease-free and overall survival were similar to sporadic cases. In the study by Park et al. (38), familial follicular cell–derived DTC patients with parent–offspring relationship were found to have a higher recurrence rate compared with sporadic cases and the second generation had even higher rates compared with the first generation. Mazeh et al. (39) found that familial DTC patients had more aggressive disease compared with sporadic cases regardless of the number of family members affected. In contrast, Robenshtok et al. (40) found that staging at diagnosis and outcomes were not different in familial DTC patients compared with sporadic DTC patients. Syndromes associated with DTC (e.g., PTEN [phosphatase and tensin homolog] hamartoma tumor syndrome [Cowden’s disease], familial adenomatous polyposis [FAP], Carney complex, multiple endocrine neoplasia [MEN] 2, Werner syndrome/progeria) in a first-degree relative, warrant screening based on various components of that syndrome (41).

  It is not possible to speculate on the impact of screening in preventing or reducing recurrence and deaths, since no interventional screening programs have ever been reported in at-risk family members. Patients with familial DTC should have a careful history and directed neck examination as a part of routine health maintenance. One should also consider thyroid cancer syndromes as noted above (41).

[A3] What is the appropriate laboratory and imaging evaluation for patients with clinically or incidentally discovered thyroid nodules?

[A4] Serum thyrotropin measurement

- **RECOMMENDATION 2**

  (A) Serum thyrotropin (TSH) should be measured during the initial evaluation of a patient with a thyroid nodule.

  (Strong recommendation, Moderate-quality evidence)

  (B) If the serum TSH is subnormal, a radionuclide (preferably ¹²³I) thyroid scan should be performed. (Strong recommendation, Moderate-quality evidence)

  (C) If the serum TSH is normal or elevated, a radionuclide scan should not be performed as the initial imaging evaluation.

  (Strong recommendation, Moderate-quality evidence)

  With the discovery of a thyroid nodule, a complete history and physical examination focusing on the thyroid gland and adjacent cervical lymph nodes should be performed. Pertinent historical factors predicting malignancy include a history of childhood head and neck radiation therapy, total body radiation for bone marrow transplantation (42), exposure to ionizing radiation from fallout in childhood or adolescence (43), familial thyroid carcinoma, or thyroid cancer syndrome (e.g., PTEN hamartoma tumor syndrome [Cowden’s disease], FAP, Carney complex, Werner syndrome/progeria, or MEN 2, a risk for medullary thyroid cancer [MTC]) in a first-degree relative, rapid nodule growth, and/or hoarseness.
Pertinent physical findings suggesting possible malignancy include vocal cord paralysis, cervical lymphadenopathy, and fixation of the nodule to surrounding tissue.

With the discovery of a thyroid nodule >1 cm in any diameter, a serum TSH level should be obtained. If the serum TSH is subnormal, a radionuclide thyroid scan should be obtained to document whether the nodule is hyperfunctioning (“hot,” i.e., tracer uptake is greater than the surrounding normal thyroid), isofunctioning (“warm,” i.e., tracer uptake is equal to the surrounding thyroid), or nonfunctioning (“cold,” i.e., has uptake less than the surrounding thyroid tissue) (44). Since hyperfunctioning nodules rarely harbor malignancy, if one is found that corresponds to the nodule in question, no cytologic evaluation is necessary. If overt or subclinical hyperthyroidism is present, additional evaluation is required. A higher serum TSH level, even within the upper part of the reference range, is associated with increased risk of malignancy in a thyroid nodule, as well as more advanced stage thyroid cancer (45,46).

**RECOMMENDATION 3**
Routine measurement of serum thyroglobulin (Tg) for initial evaluation of thyroid nodules is not recommended.

**(Strong recommendation, Moderate-quality evidence)**

Serum Tg levels can be elevated in most thyroid diseases and are an insensitive and nonspecific test for thyroid cancer (47–49).

**[A5] Serum thyroglobulin measurement**

**RECOMMENDATION 4**
The panel cannot recommend either for or against routine measurement of serum calcitonin in patients with thyroid nodules.

**(No recommendation, Insufficient evidence)**

The utility of serum calcitonin has been evaluated in a series of prospective, nonrandomized studies (50–54). These data suggest that the use of routine serum calcitonin for screening may detect C-cell hyperplasia and MTC at an earlier stage, and overall survival consequently may be improved. However, most studies relied on pentagastrin stimulation testing to increase specificity. This drug is not available in the United States, Canada, and some other countries, and there remain unresolved issues of sensitivity, specificity, assay performance, cut-offs using calcium stimulation (55), and cost effectiveness. Two retrospective studies have shown improved survival in patients diagnosed with MTC after routine calcitonin testing compared with historical controls (53,56), but they were unable to show a decreased number of MTC-related deaths. A cost-effectiveness analysis suggested that calcitonin screening would be cost effective in the United States (57). However, prevalence estimates of MTC in this analysis included patients with C-cell hyperplasia and microMTC, which have uncertain clinical significance. Based on the retrospective nature of the survival data, unresolved issues of assay performance, lack of availability of pentagastrin in North America, and potential biases in the cost-effective analysis, the task force cannot recommend for or against the routine measurement of serum calcitonin as a screening test in patients with thyroid nodules, although there was not uniform agreement on this recommendation. There was, however, agreement that serum calcitonin may be considered in the subgroup of patients in whom an elevated calcitonin may change the diagnostic or surgical approach (i.e., patients considered for less than total thyroidectomy, patients with suspicious cytology not consistent with PTC). If the unstimulated serum calcitonin determination has been obtained and the level is greater than 50–100 pg/mL, a diagnosis of MTC is common (58).

There is emerging evidence that a calcitonin measurement from a thyroid nodule fine-needle aspiration (FNA) washout may be helpful in the preoperative evaluation of patients with a modestly elevated basal serum calcitonin (20–100 pg/mL) (59).

**[A6] Serum calcitonin measurement**

**RECOMMENDATION 5**

(A) Focal $[^{18}F]$Fluorodeoxyglucose positron emission tomography ($^{18}$FDG-PET) uptake within a sonographically confirmed thyroid nodule conveys an increased risk of thyroid cancer, and FNA is recommended for those nodules ≥1 cm.

**(Strong recommendation, Moderate-quality evidence)**

B) Diffuse $^{18}$FDG-PET uptake, in conjunction with sonographic and clinical evidence of chronic lymphocytic thyroiditis, does not require further imaging or FNA.

**(Strong recommendation, Moderate-quality evidence)**

$^{18}$FDG-PET is increasingly performed during the evaluation of patients with both malignant and nonmalignant illness. While $^{18}$FDG-PET imaging is not recommended for the evaluation of patients with newly detected thyroid nodules or thyroidal illness, the incidental detection of abnormal thyroid uptake may nonetheless be encountered. Importantly, incidental $^{18}$FDG-PET uptake in the thyroid gland can be either focal or diffuse. Focal $^{18}$FDG-PET uptake in the thyroid is incidentally detected in 1%–2% of patients, while an additional 2% of patients demonstrate diffuse thyroid uptake (60–62).

Focal thyroid uptake most often corresponds to a clinically relevant thyroid nodule, and US examination is thus recommended to define thyroid anatomy. Importantly, focal $^{18}$FDG-PET uptake increases malignancy risk in an affected nodule, and therefore clinical evaluation and FNA of nodules ≥1 cm is recommended. $^{18}$FDG-PET positive thyroid nodules <1 cm that do not meet FNA criteria (see Recommendation 8) can be monitored similarly to thyroid nodules with high-risk sonographic patterns that do not meet FNA criteria (see Recommendation 24A). A recent meta-analysis confirmed that approximately one in three (~35%) $^{18}$FDG-PET positive thyroid nodules proved to be cancerous (60), with higher mean maximum standardized uptake value in malignant compared to benign nodules (6.9 vs. 4.8, $p<0.001$). In contrast, diffuse thyroid uptake most often represents benign disease corresponding to inflammatory uptake in the setting.
of Hashimoto’s disease or other diffuse thyroidal illness. However, if detected, diffuse 18FDG-PET uptake in the thyroid should also prompt sonographic examination to ensure there is no evidence of clinically relevant nodularity. Most patients with diffuse 18FDG-PET uptake demonstrate diffuse heterogeneity on sonographic examination, and no further intervention or FNA is required. It is appropriate to evaluate thyroid function in these patients.

**RECOMMENDATION 6**

Thyroid sonography with survey of the cervical lymph nodes should be performed in all patients with known or suspected thyroid nodules.

(Strong recommendation, High-quality evidence)

Diagnosis thyroid/neck US should be performed in all patients with a suspected thyroid nodule, nodular goiter, or radiographic abnormality suggesting a thyroid nodule incidentally detected on another imaging study (e.g., computed tomography [CT] or magnetic resonance imaging [MRI] or thyroidal uptake on 18FDG-PET scan) (www.aium.org/resources/guidelines/thyroid.pdf). Thyroid US can answer the following questions: Is there truly a nodule that corresponds to an identified abnormality? How large is the nodule? What is the nodule’s pattern of US imaging characteristics? Is suspicious cervical lymphadenopathy present? Is the nodule greater than 50% cystic? Is the nodule located posteriorly in the thyroid gland? These last two features might decrease the accuracy of FNA biopsy performed with palpation (63,64).

Ultrasound should evaluate the following: thyroid parenchyma (homogeneous or heterogeneous) and gland size; size, location, zand sonographic characteristics of any nodule(s); the presence or absence of any suspicious cervical lymph nodes in the central or lateral compartments. The US report should convey nodule size (in three dimensions) and location (e.g., right upper lobe) and a description of the nodule’s sonographic features including composition (solid, cystic proportion, or spongiform), echogenicity, margins, presence and type of calcifications, and shape if taller than wide, and vascularity. The pattern of sonographic features associated with a nodule confers a risk of malignancy, and combined with nodule size, guides FNA decision-making (65,66) (see Recommendation 8).

In the subset of patients with low serum TSH levels who have undergone radionuclide thyroid scintigraphy suggesting nodularity, US should also be performed to evaluate both the presence of nodules concordant with the hyperfunctioning areas on the scan, which do not require FNA, as well as other nonfunctioning nodules that meet sonographic criteria for FNA (67).

**RECOMMENDATION 7**

FNA is the procedure of choice in the evaluation of thyroid nodules, when clinically indicated.

(Strong recommendation, High-quality evidence)

FNA is the most accurate and cost-effective method for evaluating thyroid nodules. Retrospective studies have reported lower rates of both nondiagnostic and false-negative cytology from FNA procedures performed using US guidance compared to palpation (68,69). Therefore, for nodules with a higher likelihood of either a nondiagnostic cytology (>25%–50% cystic component) (64) or sampling error (difficult to palpate or posteriorly located nodules), US-guided FNA is preferred. If the diagnostic US confirms the presence of a predominantly solid nodule corresponding to what is palpated, the FNA may be performed using palpation or US guidance.

**RECOMMENDATION 8**

I. Thyroid nodule diagnostic FNA is recommended for (Fig. 2, Table 6):

(A) Nodules ≥1 cm in greatest dimension with high suspicion sonographic pattern.

(Strong recommendation, Moderate-quality evidence)

(B) Nodules ≥1 cm in greatest dimension with intermediate suspicion sonographic pattern.

(Strong recommendation, Low-quality evidence)

(C) Nodules ≥1.5 cm in greatest dimension with low suspicion sonographic pattern.

(Weak recommendation, Low-quality evidence)

II. Thyroid nodule diagnostic FNA may be considered for (Fig. 2, Table 6):

(D) Nodules ≥2 cm in greatest dimension with very low suspicion sonographic pattern (e.g., spongiform). Observation without FNA is also a reasonable option.

(Weak recommendation, Moderate-quality evidence)

III. Thyroid nodule diagnostic FNA is not required for (Fig. 2, Table 6):

(E) Nodules that do not meet the above criteria.

(Strong recommendation, Moderate-quality evidence)

(F) Nodules that are purely cystic.

(Strong recommendation, Moderate-quality evidence)

Thyroid US has been widely used to stratify the risk of malignancy in thyroid nodules, and aid decision-making about whether FNA is indicated. Studies consistently report that several US gray scale features in multivariate analyses are associated with thyroid cancer, the majority of which are PTC. These include the presence of microcalcifications, nodule hypoechochogenicity compared with the surrounding thyroid or strap muscles, irregular margins (defined as either infiltrative, microlobulated, or spiculated), and a shape taller than wide measured on a transverse view. Features with the highest specificities (median >90%) for thyroid cancer are microcalcifications, irregular margins, and tall shape,
although the sensitivities are significantly lower for any single feature (70–77). It is important to note that poorly defined margins, meaning the sonographic interface between the nodule and the surrounding thyroid parenchyma is difficult to delineate, are not equivalent to irregular margins. An irregular margin indicates the demarcation between nodule and parenchyma is clearly visible but demonstrates an irregular, infiltrative or spiculated course. Up to 55% of benign nodules are hypoechoic compared to thyroid parenchyma, making nodule hypoechogenicity less specific. In addition, subcentimeter benign nodules are more likely to be hypoechoic than larger nodules (71). Multivariable analyses confirm that the probability of cancer is higher for nodules with either microlobulated margins or microcalcifications than for hypoechoic solid nodules lacking these features (70). Macrocalcifications within a nodule, if combined with microcalcifications, confer the same malignancy risk as microcalcifications alone (70,74). However, the presence of this type of intranodal macrocalcification alone is not consistently associated with thyroid cancer (78). On the other hand, a nodule that has interrupted peripheral calcifications, in association with a soft tissue rim outside the calcification, is highly likely to be malignant, and the associated pathology may demonstrate tumor invasion in the area of disrupted calcification (79,80).

In a recent study where 98% of the cancers were PTC, intranodular vascularity did not have independent predictive value for malignancy in multivariate logistic regression model including gray-scale features (72). Two other studies and a meta-analysis with higher proportions of follicular thyroid cancer (FTC) (10%–22%) have shown that intranodular vascularity was correlated with malignancy (66,74,81). FTC exhibits other differences in sonographic features compared to PTC. These tumors are more likely to be iso- to hyperechoic, noncalcified, round (width greater than anterioposterior dimension) nodules with regular smooth margins (82). Similarly, the follicular variant of papillary cancer (FVPTC) is also more likely than conventional PTC to have this same appearance as FTC (79). Distant metastases are rarely observed arising from follicular cancers <2 cm in diameter, which therefore justifies a higher size cutoff for hyperechoic nodules (83).

The vast majority (82%–91%) of thyroid cancers are solid (70,73,75,77,84). Of 360 consecutively surgically removed thyroid cancers at the Mayo clinic, 88% were solid or minimally cystic (<5%), 9% were <50% cystic, and only 3% were more than 50% cystic (85). Therefore, FNA decision-making for partially cystic thyroid nodules must be tempered by their lower malignant risk. In addition, evidence linking sonographic features with malignancy in this subgroup of nodules is less robust, originating from univariate rather than multivariate analyses. However, an eccentric rather than concentric position of the solid component along the cyst wall, an acute rather than obtuse angle interface of the
Table 6. Sonographic Patterns, Estimated Risk of Malignancy, and Fine-Needle Aspiration Guidance for Thyroid Nodules

<table>
<thead>
<tr>
<th>Sonographic pattern</th>
<th>US features</th>
<th>Estimated risk of malignancy, %</th>
<th>FNA size cutoff (largest dimension)</th>
</tr>
</thead>
<tbody>
<tr>
<td>High suspicion</td>
<td>Solid hypoechoic nodule or solid hypoechoic component of a partially cystic nodule with one or more of the following features: irregular margins (infiltrative, microlobulated), microcalcifications, taller than wide shape, rim calcifications with small extrusive soft tissue component, evidence of ETE</td>
<td>&gt;70–90&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Recommend FNA at ≥1 cm</td>
</tr>
<tr>
<td>Intermediate suspicion</td>
<td>Hypoechoic solid nodule with smooth margins without microcalcifications, ETE, or taller than wide shape</td>
<td>10–20</td>
<td>Recommend FNA at ≥1 cm</td>
</tr>
<tr>
<td>Low suspicion</td>
<td>Isoechoic or hyperechoic solid nodule, or partially cystic nodule with eccentric solid areas, without microcalcification, irregular margin or ETE, or taller than wide shape</td>
<td>5–10</td>
<td>Recommend FNA at ≥1.5 cm</td>
</tr>
<tr>
<td>Very low suspicion</td>
<td>Spongiform or partially cystic nodules without any of the sonographic features described in low, intermediate, or high suspicion patterns</td>
<td>&lt;3</td>
<td>Consider FNA at ≥2 cm</td>
</tr>
<tr>
<td>Benign</td>
<td>Purely cystic nodules (no solid component)</td>
<td>&lt;1</td>
<td>No biopsy&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

<sup>a</sup>The estimate is derived from high volume centers, the overall risk of malignancy may be lower given the interobserver variability in sonography.

<sup>b</sup>Aspiration of the cyst may be considered for symptomatic or cosmetic drainage.

ETE, extrathyroidal extension.
solid component and cyst, and the presence of microcalcifications consistently confer a higher risk of malignancy (85–87). Other findings such as lobulated margins or increased vascularity of the solid portion are risk factors that are not as robust (86,87). However, a spongiform appearance of mixed cystic solid nodules is strongly correlated with benignity (66,70,71,88). A spongiform appearance is defined as the aggregation of multiple microcystic components in more than 50% of the volume of the nodule (71). Spongiform and other mixed cystic solid nodules may exhibit bright reflectors on US imaging, caused by colloid crystals or posterior acoustic enhancement of the back wall of a microcystic area. These may be confused with microcalcifications by less proficient sonographers, and a recent meta-analysis confirmed that operator experience is correlated with accurate evaluation of internal calcifications (66). Therefore, because of potential for misclassification, FNA may still be considered for nodules interpreted as spongiform, but with a higher size cutoff. Lastly, pure cysts, although rare (<2% of thyroid lesions), are highly likely to be benign (66,89,90).

Given the nuances in sonographic appearances of different thyroid cancer histologies, as well as the challenges posed by partially cystic nodules, some authors have suggested risk stratification based upon a constellation of sonographic features (89–91). In the absence of sonographically suspicious cervical lymph nodes, features associated with the highest risk for thyroid cancer can be used to triage smaller nodules for fine-needle biopsy, whereas nodules with sonographic appearance suggesting lower risk might be considered for fine-needle biopsy at a larger size as determined by maximal diameter (Figs. 1 and 2, Table 6). The sonographic appearance for the vast majority of thyroid nodules can be generally classified in the following categories of US patterns, which combine several individual sonographic characteristics. Since the interobserver variability in reporting individual characteristics is moderate even within controlled studies (72), the use of patterns exhibiting correlated sonographic features is more robust. Two recent studies have reported substantial interobserver correlation for identification for nodule sonographic patterns (multirater kappa statistics >0.6) (92,93).

**High suspicion [malignancy risk >70%–90% (89,90,94)].** High suspicion of malignancy is warranted with a solid hypoechoic nodule or a solid hypoechoic component in a partially cystic nodule with one or more of the following features: irregular margins (specifically defined as infiltrative, microlobulated, or spiculated), microcalcifications, taller than wide shape, disrupted rim calcifications with small extrusive hypoechoic soft tissue component, or evidence of extrathyroidal extension (Fig. 2, Table 6). A nodule demonstrating this US pattern is highly likely to be a PTC. Nodules with the high suspicion pattern and measuring ≥1 cm should undergo diagnostic fine-needle biopsy to refute or confirm malignancy. However, in the absence of evidence of extrathyroidal extension, metastatic cervical lymph nodes, or distant metastases, micropapillary thyroid cancers (<1 cm) often have an indolent course, but this may depend upon patient age (95). Although no distant metastases or deaths occurred in a recent observational series of 1235 Japanese patients with biopsy-proven PTC, tumor growth and new appearance of lymph node metastases occurred more frequently in patients younger than 40 years of age compared with those over age 60 (5.9% vs. 2.2% for size increase; 5.3% vs. 0.4% for new nodal metastases, p<0.05). Thus although a sonographically suspicious subcentimeter thyroid nodule without evidence of extrathyroidal extension or sonographically suspicious lymph nodes may be observed with close sonographic follow-up of the nodule and cervical lymph nodes, rather than pursuing immediate FNA, patient age and preference may modify decision-making (95).

**Intermediate suspicion [malignancy risk 10%–20% (89,90,94)].** Intermediate suspicion of malignancy is attached to a hypoechoic solid nodule with a smooth regular margin, but without microcalcifications, extrathyroidal extension, or taller than wide shape (Fig. 2, Table 6). This appearance has the highest sensitivity (60%–80%) for PTC, but a lower specificity than the preceding high suspicion pattern, and fine-needle biopsy should be considered for these nodules ≥1 cm to refute malignancy.

**Low suspicion [malignancy risk 5%–10% (89,90,94)].** Isoechoic or hyperechoic solid nodule, or partially cystic nodule with eccentric uniformly solid areas without microcalcifications, irregular margin or extrathyroidal extension, or taller than wide shape prompts low suspicion for malignancy (Fig. 2, Table 6). Only about 15%–20% of thyroid cancers are iso- or hyperechoic on US, and these are generally the follicular variant of PTC or FTCs (71). Fewer than 20% of these nodules are partially cystic. Therefore, these appearances are associated with a lower probability of malignancy and observation may be warranted until the size is ≥1.5 cm.

**Very low suspicion ≤3% (66,89,90,94)].** Spongiform or partially cystic nodules without any of the sonographic features described in the low, intermediate, or high suspicion patterns have a low risk of malignancy (<3%). If FNA is performed, the nodule should be at least 2 cm. Observation without FNA may also be considered for nodules ≥2 cm (Fig. 2, Table 6).

**Benign ≤1% (89,90,94)].** Purely cystic nodules are very unlikely to be malignant, and fine-needle biopsy is not indicated for diagnostic purposes (Fig. 2, Table 6). Aspiration with or without ethanol ablation may be considered as a therapeutic intervention if a cyst is large and symptomatic; cytology should be performed if aspiration is done.

Sonographic evaluation of the anterior cervical lymph node compartments (central and lateral) should be performed whenever thyroid nodules are detected. If US detects cervical lymph nodes that are sonographically suspicious for thyroid cancer (Table 7), FNA of the suspicious lymph node should be performed for cytology and washout for Tg measurement if indicated. In addition, this scenario also warrants US-guided FNA of a subcentimeter nodule that is likely to represent the primary tumor based upon sonographic features. Although there are several known clinical risk factors for thyroid cancer in patients with thyroid nodules including immobility with swallowing, pain, cough, voice change, growth, lymphadenopathy, and a history of childhood radiation therapy (either therapeutic, such as cranial radiation in childhood leukemia, or for benign conditions, such as enlarged thymus or tonsils) or familial thyroid cancer (96), these have not been incrementally included in multivariate analyses of gray-scale sonographic features and thyroid
cancer risk. However, given the higher pretest likelihood of thyroid cancer associated with these clinical risk factors, FNA can be considered at lower size cutoffs for all of the sonographic appearances described above.

Ultrasound elastography (USE) has similarly been investigated for its ability to modify thyroid cancer risk assessment among clinically relevant thyroid nodules. Elastography is a measurement of tissue stiffness. Performance requires an US machine, as well as an elastography computational module that most often must be purchased separately. An initial prospective study of 92 selected, nonrandomized patients suggested positive and negative predictive values (NPVs) near 100% (97). However, more recently, larger trials have reported substantially different results. Moon and colleagues retrospectively studied 703 thyroid nodules in comparison to gray-scale US (78). Performance of USE was inferior to that of gray-scale US assessment. The largest prospective study of 706 patients with 912 thyroid nodules was recently published by Azizi et al. (98). In this study, the positive predictive value (PPV) of USE was only 36%, comparable to that of microcalcifications. The NPV of USE was 97% in a population with cancer prevalence of 9%. Thus, while USE holds promise as a means by which to noninvasively assess cancer risk, its performance is highly variable and operator dependent. Perhaps most importantly, USE can only be effectively applied to solid nodules, thus excluding its utility for cystic or partially cystic nodules. Furthermore, to be amenable to direct pressure and determination of tissue strain, the index nodule must not overlap with other nodules in the anterioposterior plane. Obese patients, those with multinodular goiters and coalescent nodules, or patients in whom the nodule is posterior or inferior are not candidates for USE. Thus, at present, USE cannot be widely applied to all thyroid nodules in a similar fashion to gray-scale or Doppler US examination. The committee therefore believes USE (when available) may prove to be a helpful tool for preoperative risk assessment in those patients in whom accurate assessment can be performed. However, the committee cannot presently recommend its universal use or widespread adoption. Importantly, the ability to perform (or not perform) USE should not modify the recommendation for traditional gray-scale sonographic evaluation.

Finally, while most thyroid nodules meeting the preceding sonographic patterns and sizes should undergo FNA, we acknowledge that a conservative approach of active surveillance management may be appropriate as an alternative to FNA in selected patients. These may include patients with very low-risk tumors (e.g., no clinical or radiographic evidence of invasion or metastases), patients at high surgical risk, or those with a relatively short life span expectancy in whom the benefits of intervention may be unrealized.

**RECOMMENDATION 9**

Thyroid nodule FNA cytology should be reported using diagnostic groups outlined in the Bethesda System for Reporting Thyroid Cytopathology.

(Strong recommendation, Moderate-quality evidence)

To address a significant variability in the reporting of cytological findings in thyroid FNA samples, the 2007 National Cancer Institute Thyroid Fine-Needle Aspiration State of the Science Conference provided consensus recommendations known as the Bethesda System for Reporting Thyroid Cytopathology (99,100). The Bethesda system recognizes six diagnostic categories and provides an estimation of cancer risk within each category based upon literature review and expert

---

**Table 7. Ultrasound Features of Lymph Nodes Predictive of Malignant Involvement**

<table>
<thead>
<tr>
<th>Sign</th>
<th>Reported sensitivity, %</th>
<th>Reported specificity, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Microcalcifications</td>
<td>5–69</td>
<td>93–100</td>
</tr>
<tr>
<td>Cystic aspect</td>
<td>10–34</td>
<td>91–100</td>
</tr>
<tr>
<td>Peripheral vascularity</td>
<td>40–86</td>
<td>57–93</td>
</tr>
<tr>
<td>Hyperechogenicity</td>
<td>30–87</td>
<td>43–95</td>
</tr>
<tr>
<td>Round shape</td>
<td>37</td>
<td>70</td>
</tr>
</tbody>
</table>

*Adapted with permission from the European Thyroid Association guidelines for cervical ultrasound (20).*

---

**Table 8. The Bethesda System for Reporting Thyroid Cytopathology: Diagnostic Categories and Risk of Malignancy**

<table>
<thead>
<tr>
<th>Diagnostic category</th>
<th>Estimated/predicted risk of malignancy by the Bethesda system, %a</th>
<th>Actual risk of malignancy in nodules surgically excised, % median (range)b</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nondiagnostic or unsatisfactory</td>
<td>1–4</td>
<td>20 (9–32)</td>
</tr>
<tr>
<td>Benign</td>
<td>0–3</td>
<td>2.5 (1–10)</td>
</tr>
<tr>
<td>Atypia of undetermined significance or follicular lesion of undetermined significance</td>
<td>5–15</td>
<td>14 (6–48)</td>
</tr>
<tr>
<td>Follicular neoplasm or suspicious for a follicular neoplasm</td>
<td>15–30</td>
<td>25 (14–34)</td>
</tr>
<tr>
<td>Suspicious for malignancy</td>
<td>60–75</td>
<td>70 (53–97)</td>
</tr>
<tr>
<td>Malignant</td>
<td>97–99</td>
<td>99 (94–100)</td>
</tr>
</tbody>
</table>

*aAs reported in The Bethesda System by Cibas and Ali (1076).

*bBased on the meta-analysis of eight studies reported by Bongiovanni et al. (103). The risk was calculated based on the portion of nodules in each diagnostic category that underwent surgical excision and likely is not representative of the entire population, particularly of nondiagnostic and benign diagnostic categories.*
opinion (Fig. 1, Table 8). These categories are (i) nondiagnostic/unsatisfactory; (ii) benign; (iii) atypia of undetermined significance/follicular lesion of undetermined significance (AUS/FLUS); (iv) follicular neoplasm/suspicious for follicular neoplasm (FN/SFN), a category that also encompasses the diagnosis of Hurthle cell neoplasm/suspicious for Hurthle cell neoplasm; (v) suspicious for malignancy (SUSP), and (vi) malignant. Recent studies that applied the criteria and terminology of the Bethesda System to a large series of patients have shown a relatively good concordance in reporting FNA cytology, with 89%–95% of samples being satisfactory for interpretation and 55%–74% reported as definitively benign and 2%–5% as definitively malignant (101–104). The remaining samples are cytologically indeterminate, including AUS/FLUS in 2%–18% of nodules, FN in 2%–25%, and SUSP in 1%–6%. In these studies, the probability of malignancy for each Bethesda category demonstrated significant variability, but was overall compatible with the range predicted by the Bethesda System, with the exception of the AUS/FLUS diagnosis, for which the risk of malignant outcome in some studies was significantly higher than predicted (Table 8) (103,105). Recently, a blinded prospective evaluation of inter-observer concordance using Bethesda classification was performed. These data confirm an inherent limitation to the reproducibility of interpreting any cytology specimen (106). Specimens diagnosed as AUS/FLUS and SUSP were associated with the highest discordance rates. Some studies suggest that the AUS/FLUS category should be further subdivided into AUS with cytologic atypia (higher risk for malignancy) and FLUS with architectural atypia (lower risk for malignancy), but this has not yet been widely adopted (107). Nonetheless, classification using the Bethesda system has proven highly beneficial, allowing practitioners to speak with certainty when the cystic component is <50% (64,112,117). Nodules with larger cystic portion have a higher chance to yield nondiagnostic samples on the initial and repeated FNA.

Most nodules with a nondiagnostic cytology interpretation are benign. In large series of patients classified based on the Bethesda System, nondiagnostic samples constituted 2%–16% of all FNA samples, of which 7%–26% were eventually resected (101–103). The frequency of malignancy among all initially nondiagnostic samples was 2%–4% and among those nondiagnostic samples that were eventually resected was 9%–32%.

Sonographic features are also useful for identifying which nodules with repeat nondiagnostic FNA cytology results are more likely to be malignant. Of 104 nodules with two nondiagnostic cytology results, thyroid cancer was found in 25% of those with microcalcifications, irregular margins, a taller than wide shape, or hypoechogenicity, but in only 4% lacking these features (118).

In some studies, the use of thyroid core-needle biopsy (119) and molecular testing for BRAF (120,121) or a panel of mutations (122) helped to facilitate appropriate management of these patients, although the full clinical impact of these approaches for nodules with nondiagnostic cytology remains unknown. Some studies have found that core biopsy offers a higher adequacy rate, but may be less sensitive for the detection of papillary cancer (123,124). Mutational testing may be informative in samples considered inadequate by qualitative criteria (i.e., due to poor preparation or poor staining of cells) but is unlikely to be contributory in samples with insufficient quantity of cells.

### [A12] Nondiagnostic cytology

**RECOMMENDATION 10**

(A) For a nodule with an initial nondiagnostic cytology result, FNA should be repeated with US guidance and, if available, on-site cytologic evaluation

*(Strong recommendation, Moderate-quality evidence)*

(B) Repeatedly nondiagnostic nodules without a high suspicion sonographic pattern require close observation or surgical excision for histopathologic diagnosis

*(Weak recommendation, Low-quality evidence)*

(C) Surgery should be considered for histopathologic diagnosis if the cytologically nondiagnostic nodule has a high suspicion sonographic pattern, growth of the nodule (>20% in two dimensions) is detected during US surveillance, or clinical risk factors for malignancy are present

*(Weak recommendation, Low-quality evidence)*

Nondiagnostic or unsatisfactory FNA biopsies are those that fail to meet the established quantitative or qualitative criteria for cytologic adequacy (i.e., the presence of at least six groups of well-visualized follicular cells, each group containing at least 10 well-preserved epithelial cells, preferably on a single slide) (99,108). Although an FNA specimen found to have abundant colloid and few epithelial cells may be considered nondiagnostic by the above criteria, this is also likely a benign biopsy. After an initial nondiagnostic cytology result, repeat FNA with US guidance and, if available, on-site cytologic evaluation, will substantially increase the rate of specimen adequacy (109–113). It has been suggested that repeat FNA should be performed no sooner than 3 months after the initial FNA to prevent false-positive interpretation due to biopsy-induced reactive/reparative changes (114). Two recent studies have questioned the necessity for a 3-month waiting period after the first FNA because they did not find a correlation between the diagnostic yield/accuracy of repeated FNA and the waiting time between the procedures (115,116). A 3-month waiting period after a nondiagnostic biopsy is likely not necessary. If clinical and US features are suspicious for malignancy, a shorter waiting period may be appropriate. Repeat FNA with US guidance will yield a diagnostic cytology specimen in 60%–80% of nodules, particularly when the cystic component is <50% (64,112,117). Nodules with larger cystic portion have a higher chance to yield nondiagnostic samples on the initial and repeated FNA.

Accurate FNA cytology diagnosis depends upon a number of factors including the skill of the operator, FNA technique, specimen preparation, and cytology interpretation.
Ultrasound-guided FNA with real-time visualization of needle placement in the target nodule decreases the false-negative rate of a benign cytology diagnosis (68,69,126,128). Although prospective studies are lacking, malignancy rates of only 1%–2% have been reported in large retrospective series that analyzed the utility of systematic repeat FNA in nodules with prior benign cytology results (129–133). A pooled analysis of 12 studies by Tee et al. (134) showed that of 4055 patients with benign cytology who underwent surgery, the rate of malignancy was 3.2%.

Studies have also attempted to correlate nodule size with accuracy of FNA cytology. Several surgical series have reported higher malignancy rates in nodules >3–4 cm, but these suffer from both selection bias (only a subset of patients underwent preoperative FNA) and potential sampling error (FNA performed by palpation) (135,136). A recent study evaluated the accuracy of FNA and US features in patients with thyroid nodules ≥4 cm (137). This was a single-center study in which the practice is to offer thyroidecomy or lobectomy to all patients with nodules ≥4 cm. The investigators identified thyroid cancer in 22% of 382 nodules. A subset of thyroid nodules underwent preoperative FNA, and of the 125 cytologically benign nodules, 10.4% were malignant on final histopathology. The investigators further showed that individual US characteristics were not predictive for malignancy, although they did not look at sonographic patterns. Two other recent reports of consecutive US-guided FNA evaluations in over 1400 nodules ≥3 cm with initial benign cytology followed for a mean of 3 years confirmed a lower false-negative rate of <1.5% (138,139). Interestingly, in both these studies 66% of the missed cancers were found in nodules with high suspicion sonographic pattern, despite initial benign cytology. A recent retrospective study analyzed the long-term follow-up of 2010 cytologically benign nodules from 1369 patients. Over a mean follow-up of 8.5 years 18 false-negative malignancies were detected. However, no deaths attributable to thyroid cancer were identified in this cohort. These data confirm that an initially benign FNA confers negligible mortality risk during long-term follow-up despite a low but real risk of false negatives in this cytologic category (140). Based on the evidence, it is still unclear if patients with thyroid nodules ≥4 cm and benign cytology carry a higher risk of malignancy and should be managed differently than those with smaller nodules.

Follow-up for patients with benign cytology is discussed in section [A24] and Recommendation 23.

[A14] Malignant cytology

RECOMMENDATION 12

If a cytology result is diagnostic for primary thyroid malignancy, surgery is generally recommended.

(Strong recommendation, Moderate-quality evidence)

A cytology diagnostic for a primary thyroid malignancy will almost always lead to thyroid surgery. However, an active surveillance management approach can be considered as an alternative to immediate surgery in

(A) patients with very low risk tumors (e.g., papillary microcarcinomas without clinically evident metastases or local invasion, and no convincing cytologic evidence of aggressive disease),

(B) patients at high surgical risk because of comorbid conditions,

(C) patients expected to have a relatively short remaining life span (e.g., serious cardiopulmonary disease, other malignancies, very advanced age), or

(D) patients with concurrent medical or surgical issues that need to be addressed prior to thyroid surgery.

Following thyroid surgery for papillary thyroid microcarcinoma (PTMC), defined as a tumor 1 cm or less in size, disease-specific mortality rates have been reported to be <1%, loco-regional recurrence rates are 2%–6%, and distant recurrence rates are 1%–2% (141,142). It is quite likely that these excellent outcomes are more related to the indolent nature of the disease rather than to the effectiveness of treatment, since two prospective clinical studies of active surveillance from Japan reported similar clinical outcomes in 1465 patients with biopsy-proven PTMC that were not surgically removed and were followed for up to 15 years (average 5–6 years, range 1–17 years) (95,143). In the study by Ito et al. (95), observation was offered to 1235 patients with PTMC that did not have (i) location adjacent to the trachea or on the dorsal surface of the lobe close to the recurrent laryngeal nerve, (ii) FNA findings suggestive of high-grade malignancy; (iii) presence of regional lymph node metastases; or (iv) signs of progression during follow-up. Of those, most patients showed stable tumor size on average follow-up of 60 months (range 18–227 months), whereas 5% showed tumor enlargement (≥3 mm) by US on 5-year follow-up, and 8% on 10-year follow-up. Furthermore, 1.7% and 3.8% of patients at 5-year and 10-year follow-up showed evidence for lymph node metastases. Of 1235 patients, 191 underwent surgical treatment after observation, including those with tumor enlargement and new lymph node metastases. These patients have been followed an average of 75 months (range 1–246 months) after the surgical intervention. Only one of the patients treated with surgery after observation developed tumor recurrence. In the study by Sugitani et al. (143), 230 patients with asymptomatic PTMC were followed for 5 years on average. Of those patients, tumor size enlargement was observed in 7%, and 1% developed apparent lymph node metastasis. Seven percent of patients underwent surgery after 1–12 years of follow-up, and no recurrences were identified in those on limited follow-up, suggesting that the delayed surgery did not affect the outcome. A more recent study by Ito and colleagues followed 1235 patients with PTMC under active surveillance for an average of 60 months (95). Only 43 patients (3.5%) had clinical progression of disease by their stated criteria (tumor growing to >12 mm or appearance of new lymph node metastases). Interestingly, the younger patients (<40 years old) had an 8.9% rate of clinical progression, while those 40–60 years old had a 3.5% rate of progression and those >60 years old had the lowest rate of clinical progression (1.6%).

Despite the evidence that cautious observation is a safe and effective alternative to immediate surgical resection, very few PTMC patients outside of those two centers in Japan are given the option of an active surveillance approach. This is in part due to reports in the literature of a small percentage of patients with PTMC presenting with clinically significant regional or distant metastases (141,142,144). Unfortunately, no clinical features (145–151) can reliably differentiate the relatively small number
of PTMC patients destined to develop clinically significant progression from the larger population of people that harbor indolent PTMCs that will not cause significant disease.

Similarly, well-known thyroid cancer oncogenes, such as BRAF, when taken in isolation, are not able to specifically identify the microcarcinomas that will progress and spread outside of the thyroid. The prevalence of BRAF V600E mutations in PTMC with lymph node metastases and tumor recurrence is higher than PTMC without LN metastases or recurrence, and in some studies the presence of a BRAF mutation was associated with lymph node metastasis from PTMC on multivariate analysis (150,152,153). These studies showed that although the presence of a BRAF V600E mutation identifies 65%–77% of patients with PTMC that develop lymph node metastases, the BRAF status taken in isolation has a low PPV for detecting PTMC with extrathyroidal spread and therefore has a limited role for guiding patient management. However, recent data suggest that specific molecular profiles, such as the coexistence of BRAF with other oncogenic mutations (such as PIK3CA, AKTI), TERT promoter, or TP53 mutations may serve as more specific markers of less favorable outcome of PTC. Therefore, it is likely that finding of one of these genetic profiles in a small tumor would suggest that it represents an early stage of a clinically relevant PTC (154–157). Future studies are expected to establish the impact of molecular profiling involving multiple mutations or other genetic alterations on clinical management of patients with PTMC.

**[A15]** Indeterminate cytology (AUS/FLUS, FN, SUSP)*

**[A16]** What are the principles of the molecular testing of FNA samples?

Molecular markers may be classified according to intended use; that is, diagnostic (classification of a disease state), prognostic, or predictive purposes (providing information on the estimated probability of beneficial therapy or harm of a specific therapy) (158). Furthermore, companion use of predictive molecular markers involves the identification of patient subgroups in which a therapeutic intervention is proven to be either beneficial or harmful, with intended implications for appropriate clinical stratification of therapies (158). Validation studies of molecular marker tests may include evaluation of (a) analytic validity (including test accuracy and reproducibility in ascertaining the molecular event), (b) clinical validity (the performance of the test in distinguishing different groups of patients, based on biology or expected disease outcome, including measures of sensitivity and specificity or predictive values), and (c) clinical utility (examination of the test’s ability to improve outcomes, with direct clinical decision-making implications) (158). Furthermore, an NCCN Tumor Marker Task Force has indicated that the clinical utility of a molecular test should be founded in strong evidence proving that use of the marker “improves patient outcomes sufficiently to justify its incorporation into routine clinical practice.”†

The principal proposed use of molecular markers in indeterminate thyroid FNA specimens is diagnostic (ruling out or in the presence of thyroid malignancy), with the implication of a companion use to inform decision-making on primary surgical treatment (i.e., the decision to perform surgery and if so, the extent of surgery). However, the focus of this section is restricted to the clinical validity of molecular testing of indeterminate FNA specimens. It is important to note that long-term outcome data on companion use of molecular marker status to guide therapeutic decision-making is currently lacking, and therefore we do not know if implementation of molecular marker use in routine clinical practice would result in a significant overall benefit in health outcomes in patients with thyroid nodules. Surgical decision-making on indeterminate FNA specimens is reviewed in another section of these guidelines, with some reference to molecular marker testing (if performed). As summarized in a Disease State Commentary from the AACE Thyroid Scientific Committee and a consensus statement from the ATA Surgical Affairs Committee, use of molecular marker testing on indeterminate FNA specimens should not be intended to replace other sources of information or clinical judgment (159,160). The pretest probability of malignancy (based on clinical risk factors, cytology, US findings), feasibility considerations, and patient preferences are some additional factors that need to be considered in decision-making related to molecular marker testing of FNA specimens. Because this is a rapidly evolving field of investigation, it will be important to perform interval evaluations of the published evidence to ensure that recommendations remain contemporary.

A number of molecular approaches have been studied in the clinical setting of indeterminate FNA cytologic interpretation (161). One could surmise that an ideal “rule-in” test would have a PPV for histopathologically proven malignancy similar to a malignant cytologic diagnosis (98.6%), and an ideal “rule-out” test would have a NPV similar to a benign cytologic diagnosis (96.3%) (predictive value estimates based on a recent meta-analysis of performance of the Bethesda system) (103), and these would hold true with a reasonable degree of precision and reproducibility.

### RECOMMENDATION 13

If molecular testing is being considered, patients should be counseled regarding the potential benefits and limitations of testing and about the possible uncertainties in the therapeutic and long-term clinical implications of results.

**(Strong recommendation, Low-quality evidence)**

The largest studies of preoperative molecular markers in patients with indeterminate FNA cytology have respectively evaluated a seven-gene panel of genetic mutations and rearrangements (BRAF, RAS, RET/PTC, PAX8/PPARγ) (162), a gene expression classifier (167 GEC; mRNA expression of 167 genes) (163), and galectin-3 immunohistochemistry (cell blocks) (164). These respective studies have been subject to various degrees of blinding of outcome assessment (162–164).

Mutational testing has been proposed for use as a rule-in test because of relatively high reported specificity (86%–100%)
and PPV (84%–100%) (105,122,162,165–168). Although BRAFV600E single mutation testing has been estimated to have a specificity of approximately 99% (pooled data from 1117 nodules with histopathologic confirmation from multiple studies), the sensitivity has been deemed to be too low to reliably rule out the presence of malignancy (169). Therefore, mutational panels have been expanded to include multiple mutations/translocations including BRAF, NRAS, HRAS, and KRAS point mutations, as well as RET/PTC1 and RET/PTC3, with or without PAX8/PPARγ rearrangements (105,122,162,165–168) and other gene rearrangements (170). In indeterminate cytology thyroid nodules, the sensitivity of the seven-gene mutational panel testing is variable, with reports ranging from 44% to 100% (162,165,167). The reported variability in sensitivity of mutational analysis with the seven-gene panel in indeterminate nodules suggests that traditional limited mutation panels may not reliably rule out malignancy with a negative test in this population. Next-generation sequencing of an expanded panel of point mutations, single base insertions/deletions (indels), and gene rearrangements has been reported to have a sensitivity of 90% for FN/SFN FNA cytology specimens from a single-center study (170). A limitation was that the pathologists evaluating the gold standard surgical pathology specimens were aware of the results of earlier generation molecular tests previously conducted on the FNAs (170). It is not known to what extent differences in techniques used to perform mutational testing by various groups (162,167,170) may affect test performance, and direct, head-to-head comparisons of these tests within the same population are lacking. The currently available seven-gene mutational panels have been proposed to be most useful when surgery is favored. However, this is based on the assumption that the surgical approach would be altered with a positive test, and long-term outcome data proving the overall benefit of this therapeutic strategy are needed. It is important to acknowledge that algorithms employing seven-gene mutational testing as a means to inform decision-making on extent of primary thyroid surgery (i.e., lobectomy or total thyroidectomy) (162) were developed at a time when the ATA guidelines favored total thyroidectomy for most PTCs >1 cm in diameter (25). However, this does not reflect recommendations in these guidelines (see Recommendation 35 on surgical management of malignant cytology nodules). Furthermore, long-term outcome data from a strategy of using molecular markers in indeterminate FNA specimens to stratify surgical approach are currently lacking.

Use of a 167 GEC has been proposed as a rule-out test due to relatively high sensitivity (92%) and NPV (93%), as reported in a prospective multicenter study (163). The relatively low specificity of the 167 GEC test (mean values 48%–53% in indeterminate nodules subject to histopathologic confirmation) suggests that the test cannot definitively rule in malignancy in indeterminate nodules. In a retrospective analysis of 167 GEC results from five institutions, Alexander et al. (171) reported that the prevalence of 167 GEC benign readings by institution varied up to 29%, which was not statistically significant. The distribution of recruitment from each of the five study sites was highly variable (total n = 339 nodules), with two sites contributing only 30 and 37 patients, and the other three sites accounting for the majority of the study population. The prevalence of malignancy confirmed by histopathology in 167 GEC “suspicious” nodules ranged from 33% to 80% in the 48 nodules in the AUS/FLUS group across institutions (no 95% confidence interval [95% CI] reported, p = 0.11), and from 33% to 67% in the 65 nodules in the FN cytology group (no 95% CI reported, p = 0.87) (171). For the 174 patients with 167 GEC “benign” readings, four patients were advised to undergo surgery (2%), and 41% (71 patients) of this group had short documented follow-up for a mean of 8.5 months (median 8 months, range 1–24 months); ultimately, 6% of patients in this group (11/174) had surgery, with one histopathologically confirmed malignancy (171). The reproducibility of 167 GEC NPV measures in different populations of patients with indeterminate cytology thyroid nodules has recently been questioned in three smaller, independent, unblinded studies (172–174). None of these studies reported any degree of blinding. Furthermore, 95% CI of predictive estimates in indeterminate cytology nodules were not reported in two single-center studies (172,173), and one multi-center study was reported to be a retrospective analysis (174), making it difficult to interpret the findings. However, such data highlight the need for additional independent research examining the reproducibility of diagnostic efficacy of the 167 GEC in more institutions, and the importance of reporting precision estimates for diagnostic accuracy measures. Furthermore, as in the case of other molecular-based diagnostic tests in the field, long-term outcome data from clinical utility studies are needed to inform potential future clinical practice implications of the 167 GEC.

Immunohistochemical stains such as galectin-3 and HBME-1 have been examined in multiple studies of histologically confirmed thyroid FNA samples with indeterminate cytology, with reports of relatively high rates of specificity, but low sensitivity, for cancer detection (164,175,176). Two of these studies were reported to be prospective (164,175), but only one of the studies reported any degree of blinding (blind central histopathologic review) (164). Immuno-histochemical stains require the availability of a cell block to perform the staining. Additional diagnostic molecular marker strategies are also under development. Specifically, mRNA markers (177–179), as well as miRNA markers (180–185), have shown initial diagnostic utility in FNA samples with indeterminate cytological diagnoses, but they have not been thoroughly validated. A recent study combining seven-gene mutational testing with expression of a set of 10 miRNA genes on preoperative FNA sampling from 109 patients with AUS/FLUS or FN cytology, showed 89% sensitivity, 85% specificity, with a 73% PPV and 94% NPV on this group with a 32% prevalence of malignancy (186). Finally, peripheral blood TSH receptor mRNA assay has been reported to have a 90% PPV and 39% NPV in FNA-based assessment of thyroid nodules with atypical or suspicious cytology in a single-center, prospective validation study (187).

In summary, there is currently no single optimal molecular test that can definitively rule in or rule out malignancy in all cases of indeterminate cytology, and long-term outcome data proving clinical utility are needed.

**RECOMMENDATION 14**

If intended for clinical use, molecular testing should be performed in Clinical Laboratory Improvement Amendments/College of American Pathologists (CLIA/CAP)-certified molecular laboratories, or the international equivalent,
because reported quality assurance practices may be superior compared to other settings.

(Strong recommendation, Low-quality evidence)

Many molecular marker tests are available in hospital-based molecular pathology laboratories and in reference laboratories. Importantly, all molecular marker tests intended for clinical use should be performed only in CLIA/CAP-certified molecular laboratories after appropriate analytical and clinical validation of all assays in each laboratory (158). In a survey of American molecular genetic testing laboratory directors, laboratory process quality assurance score (for multiple relevant domains) was associated with the presence of CLIA certification (188). In a large, international survey of medical genetic testing laboratory directors, accreditation of the laboratory was associated with a higher quality assurance index score (189).

[A17] AUS/FLUS cytology

■ RECOMMENDATION 15

(A) For nodules with AUS/FLUS cytology, after consideration of worrisome clinical and sonographic features, investigations such as repeat FNA or molecular testing may be used to supplement malignancy risk assessment in lieu of proceeding directly with a strategy of either surveillance or diagnostic surgery. Informed patient preference and feasibility should be considered in clinical decision-making.

(Weak recommendation, Moderate-quality evidence)

(B) If repeat FNA cytology, molecular testing, or both are not performed or inconclusive, either surveillance or diagnostic surgical excision may be performed for an AUS/FLUS thyroid nodule, depending on clinical risk factors, sonographic pattern, and patient preference.

(Strong recommendation, Low-quality evidence)

Based on the Bethesda System, this diagnostic category is reserved for specimens that contain cells with architectural and/or nuclear atypia that is more pronounced than expected for benign changes but not sufficient to be placed in one of the higher risk diagnostic categories (99,190). Although this diagnostic category has been recommended for limited use and has an expected frequency in the range of 7%, recent reports based on the Bethesda System have found this cytologic diagnosis to be used in 1%–27% of all thyroid FNA samples (105,191). In studies that utilized the criteria established by the Bethesda System, the risk of cancer for patients with AUS/FLUS nodules who underwent surgery was 6%–48%, with a mean risk of 16% (191).

A second opinion review of the cytopathology slides by a high-volume cytopathologist may be considered for patients with AUS/FLUS cytology. There is some evidence that this approach may reclassify many of these patients into the benign and nondiagnostic categories (106,192). Furthermore, the overall diagnostic accuracy may be improved. Unfortunately, there is a relatively high intra-observer variability in this difficult diagnostic category (106).

A repeat FNA yields a more definitive cytologic diagnosis in many cases, whereas 10%–30% of nodules are repeatedly AUS/FLUS (104,193–195). The rate of malignancy on surgical follow-up has been shown to be similar for patients with a single AUS/FLUS diagnosis (37/90, 41%), two successive AUS/FLUS diagnoses (22/51, 43%), and patients with a benign cytologic interpretation following the initial AUS/FLUS diagnosis (27/29%) (196). Use of thyroid core-needle biopsy was reported by some to be more informative than repeated FNA for sampling nodules that were AUS/FLUS on initial FNA (119), and it is reasonably well-tolerated (197).

The refinement of risk stratification of nodules with AUS/FLUS cytology using molecular testing has been examined in multiple studies. The interpretation of molecular testing is complex, however, and its utility is strongly influenced by the prevalence of cancer in the tested population of nodules. Only two molecular tests have been separately evaluated and validated for the individual AUS/FLUS, FN, and SUSP categories. Mutational testing for BRAF in AUS/FLUS samples has high specificity for cancer, but low sensitivity (198,199). Testing for a panel of mutations (BRAF, NRAS, HRAS, KRAS, RET/PTC1, RET/PTC3, PAX8/PPARγ) offers a significantly higher sensitivity of 63%–80% (162,165). In the largest prospective study to date of nodules with AUS/FLUS cytology (653 consecutive nodules, of which 247 had surgical follow-up) from a single institution, detection of any of these mutations using RT-PCR with fluorescent melting curve analysis was reported to convey an 88% risk of cancer among nodules with surgical follow-up; 63% of cancers on final histopathology were identified with a positive mutation preoperatively (22 of 35), and 94% of nodules that were negative on mutation analysis had a benign final histopathology (209 of 222) (162). Positive testing for BRAF, RET/PTC or PAX8/PPARγ was specific for a malignant outcome in 100% of cases, whereas RAS mutations had an 84% risk of cancer and a 16% chance of benign follicular adenoma. A recent study utilizing RT-PCR with liquid bead array flow cytometry with the seven-gene mutation panel reported on 11 AUS/FLUS cytology nodules that had malignant histopathologic confirmation, of which seven had a negative molecular test and four had a positive test (167). There were also 11 AUS/FLUS cytology nodules, which were benign on histopathologic evaluation; 9 of 11 had a negative molecular test result and 2 of 11 had a positive result. Interpretation of results from the AUS/FLUS subgroup is limited by the small reported sample size (167).

Molecular testing using the 167 GEC has been studied for its diagnostic use in nodules with AUS/FLUS cytology. A multi-institutional study of 129 FNA samples with AUS/FLUS cytology and surgical follow-up reported a 90% sensitivity [95% CI 74%–98%] and 95% NPV [95% CI 85%–99%], but only a 53% specificity [95% CI 43%–63%] and 38% PPV [95% CI 27%–50%] for cancer (163). Although the specificity of the 167 GEC was low (53% [95% CI 43%–63%]), the negative test result was reported to decrease the risk estimate of malignancy in AUS/FLUS nodules in this study from 24% to 5%. This observation has led to a clinical extrapolation suggesting that nodules that have a negative 167 GEC test results may be followed without surgery (163). In a recent single-center retrospective study including 68 cases of AUS/FLUS nodules, 16 AUS/FLUS cases were reported to have a 167 GEC suspicious result, and the PPV was 61% (11 of 18) for those with surgical pathology confirmation (200). There were insufficient data to confirm the NPV of the 167 GEC test for AUS/FLUS cytology, since the vast majority of patients with a benign 167 GEC test did not undergo surgery, and no long-term follow-up of such cases was reported.
In three other recent studies, there were insufficient data for analysis in the AUS/FLUS subgroup to draw any meaningful conclusion on 167 GEC test performance in that subgroup (172–174). In addition, published follow-up for the 167 GEC is currently limited to a mean of 8.5 months in a subgroup of 71 patients (171), hence long-term outcome data are needed to ensure durability of benign 167 GEC findings with correlation to clinical and histologic outcomes.

Several recent studies (201–205) have examined the feasibility of using sonographic features to estimate risk of malignancy in nodules with AUS/FLUS cytology. The PPV of suspicious sonographic features has been estimated to range from 60% to 100% depending on the pretest probability of malignancy in AUS/FLUS cytology and the specific sono-

The findings of these studies must be interpreted in the context of each study’s overall risk of malignancy for this cytology classification because of its effect on the PPVs obtained by subsequent application of sonographic features to cytologically AUS/FLUS nodules. From the four Korean studies (overall malignancy rate 40%–55%), the reported cancer risk in AUS/FLUS nodules with the high suspicion sonographic pattern is 90%–100% (201–204), and the presence of even one suspicious US feature (irregular margins, taller than wide shape, marked hypoechogenicity or microcalcifications) increases the cancer risk to 60%–90%. However, when the reported cancer rate in AUS/FLUS nodules is lower (e.g., 23% in a study from Brazil (205)), the high suspicion sonographic pattern still raises the risk of malignancy, but the PPV is lower at 70%. Nonetheless, the incidence of cancer in AUS/FLUS nodules with either the high suspicion pattern US or just one suspicious US feature is significantly higher than that generally accepted for this cytology category. In a secondary analysis of a retrospective single-center study of indeterminate FNA specimens subject to 167 GEC testing, Lasra et al. studied whether re-

The refinement of risk stratification of nodules with FN/ SFN/Hürthle cell neoplasm cytology has been examined using ancillary molecular testing. Testing for a seven-gene panel of mutations (including BRAF, RAS, RET/PTC, and PPARγ) in nodules with follicular or Hürthle cell neoplasm or SFN has been reported to be associated with a sensitivity of 57%–75%, specificity of 97%–100%, PPV of 87%–100%, and NPV of 79%–86% (162,165). Many of these benign tumors are follicular adenomas driven by oncogenic RAS mutation with uncertain malignant potential (208). Nodules lacking all of these mutations still have a substantial cancer risk, which is due to the presence of a subset of tumors that lack any of the mutations tested by this seven-gene panel (162). Expansion of the current panels to include additional mutations and gene rearrangements with this next-generation sequencing assay was associated with a sensitivity of 90% [95% CI 80%–99%], specificity of 93% [95% CI 88%–98%], PPV of 83% [95% CI 72%–95%], and NPV of 96% [95% CI 92%–100%] in a recent single-center study of 143 consecutive FN/SFN FNA specimens with known surgical outcomes. In this study (170), retrospective (n=91) and prospective data (n=52) were combined. A limitation was that the pathologists reviewing the surgical specimens were aware of earlier generation molecular marker seven-gene panel test results, although they were blinded to results of the next-

This diagnostic category of the Bethesda System is used for cellular aspirates (i) comprised of follicular cells arranged in an altered architectural pattern characterized by cell crowding and/or microfollicle formation and lacking nuclear features of papillary carcinoma or (ii) comprised almost exclu-

(A) Diagnostic surgical excision is the long-established standard of care for the management of FN/SFN cytology nodules. However, after consideration of clinical and sonographic features, molecular testing may be used to supplement malignancy risk assessment data in lieu of proceeding directly with surgery. Informed patient preference and feasibility should be considered in clinical decision-making.

(B) If molecular testing is either not performed or inconclu-

(Weak recommendation, Moderate-quality evidence)

(Strong recommendation, Low-quality evidence)
15% (4 of 27), and NPV 75% (three of four) (173). In a single-center retrospective study including 64 nodules subjected to 167 GEC testing and a cytology read as FN/FN with oncocytic features, the PPV for a suspicious GEC result was 37% (11 of 30), although the PPV was significantly higher in the FN group (53%) compared with the FN with oncocytic features group (15%) (200). There were insufficient numbers of patients with benign 167 GEC results who underwent surgery to draw conclusions on NPV; moreover, no long-term follow-up data were reported (200). The relatively small number of indeterminate nodules is an important limitation. Furthermore, precision estimates (95% CIs) for the diagnostic accuracy measures were not reported but could be assumed to be relatively wide given the small sample size (173,200).

[A19] Suspicious for malignancy cytology

■ RECOMMENDATION 17

(A) If the cytology is reported as suspicious for papillary carcinoma (SUSP), surgical management should be similar to that of malignant cytology, depending on clinical risk factors, sonographic features, patient preference, and possibly results of mutational testing (if performed).

(Strong recommendation, Low-quality evidence)

(B) After consideration of clinical and sonographic features, mutational testing for BRAF or the seven-gene mutation marker panel (BRAF, RAS, RET/PTC, PAX8/PPARγ) may be considered in nodules with SUSP cytology if such data would be expected to alter surgical decision-making.

(Weak recommendation, Moderate-quality evidence)

This diagnostic category of the Bethesda System is reserved for aspirates with cytologic features that raise a strong suspicion for malignancy (mainly for PTC) but are not sufficient for a conclusive diagnosis (99,209). This is the highest risk category of indeterminate cytology in the Bethesda System, with an estimated cancer risk of 60%–75% (209). Studies that utilize the Bethesda System have reported this cytologic diagnosis in 1%–6% of thyroid FNAs and found malignancy after surgery in 53%–87% (mean, 75%) of these nodules (191). Due to the high risk of cancer, the diagnosis of suspicious for papillary carcinoma is an indication for surgery.

Mutational testing has been proposed to refine risk prior to surgery, assuming that surgical management would change based on a positive test result. BRAF mutations have been reported to confer close to 100% probability of malignancy (162,198,210,211). Testing for BRAF mutations in nodules suspicious for malignancy has been reported to have a sensitivity of 36% (10 of 28) and specificity of 100% (24 of 24) in a single-center retrospective study (162). In another single-center retrospective study in which FNA slides deemed to be suspicious for PTC were tested after surgery, the sensitivity of BRAF testing for PTC was 45.5% (15 of 33), and specificity was 85.7% (12 of 14) (212). Testing for a seven-gene panel of mutations (including BRAF, RAS, RET/PTC, with or without PAX8/PPARγ) in nodules with cytologic suspicious for malignancy is associated with a sensitivity of 50%–68%, specificity of 86%–96%, PPV of 80%–95%, and NPV of 72%–75% in respective single-center studies (162,165,168). Molecular testing using the 167 GEC has a PPV that is similar to cytology alone (76%) and a NPV of 85% (163), and it is therefore not indicated in patients with this cytologic diagnosis.

[A20] What is the utility of 18FDG-PET scanning to predict malignant or benign disease when FNA cytology is indeterminate (AUS/FLUS, FN, SUSP)?

■ RECOMMENDATION 18

18FDG-PET imaging is not routinely recommended for the evaluation of thyroid nodules with indeterminate cytology.

(Weak recommendation, Moderate-quality evidence)

Eight studies have been performed and are the subject of two meta-analyses (213–222). While early data suggested a high NPV for 18FDG-PET in this setting, most studies failed to use the Bethesda System for Reporting Thyroid Cytopathology and included numerous small nodules <1 cm in diameter (221). A recent meta-analysis included seven studies, of which five were prospective (222). The cancer prevalence was 26% inclusive of all combined data, confirming a typical study cohort. Sensitivity and specificity of 18FDG-PET were 89% and 55%, respectively, resulting in a 41% PPV and 93% NPV, which is similar to the performance of the 167 GEC. Vriens et al. (223) performed a cost-effectiveness analysis using 18FDG-PET performance data from their own meta-analysis and 2012 reimbursement rates of the Dutch system. They showed that 18FDG-PET was more cost effective than surgery, the 167 GEC, or mutational testing. A recent prospective analysis of 56 nodules with indeterminate FNA cytology used both 18FDG-PET and thyroid US to further evaluate the nodules (220). In a multivariate analysis, the authors demonstrated no additional diagnostic benefit or improved risk assessment when adding 18FDG-PET to that already obtained from thyroid US, bringing into question the incremental benefit of PET imaging in patients with cytologically indeterminate thyroid nodules.

[A21] What is the appropriate operation for cytologically indeterminate thyroid nodules?

■ RECOMMENDATION 19

When surgery is considered for patients with a solitary, cytologically indeterminate nodule, thyroid lobectomy is the recommended initial surgical approach. This approach may be modified based on clinical or sonographic characteristics, patient preference, and/or molecular testing when performed (see Recommendations 13–16).

(Strong recommendation, Moderate-quality evidence)

■ RECOMMENDATION 20

(A) Because of increased risk for malignancy, total thyroideectomy may be preferred in patients with indeterminate nodules that are cytologically suspicious for malignancy, positive for known mutations specific for carcinoma, sonographically suspicious, or large (>4 cm), or in patients with familial thyroid carcinoma or history of radiation exposure, if completion thyroideectomy would be
recommended based on the indeterminate nodule being malignant following lobectomy.

**(Strong recommendation, Moderate-quality evidence)**

(B) Patients with indeterminate nodules who have bilateral nodular disease, those with significant medical comorbidities, or those who prefer to undergo bilateral thyroideectomy to avoid the possibility of requiring a future surgery on the contralateral lobe, may undergo total or near-total thyroideectomy, assuming completion thyroideectomy would be recommended if the indeterminate nodule proved malignant following lobectomy.

**(Weak recommendation, Low-quality evidence)**

The primary goal of thyroid surgery for a thyroid nodule that is cytologically indeterminate (AUS/FLUS or FN or SUSP) is to establish a histological diagnosis and definitive removal, while reducing the risks associated with remedial surgery in the previously operated field if the nodule proves to be malignant. Surgical options to address the nodule should be limited to thyroideectomy (hemithyroidectomy) with or without isthmusectomy, near-total thyroideectomy (removal of all grossly visible thyroid tissue, leaving only a small amount [<1 g] of tissue adjacent to the recurrent laryngeal nerve near the ligament of Berry), or total thyroideectomy (removal of all grossly visible thyroid tissue). Removal of the nodule alone, partial thyroideectomy, and subtotal thyroideectomy, leaving >1 g of tissue with the posterior capsule on the uninvolved side, are inappropriate operations for possible thyroid cancer (224).

Decisions regarding the extent of surgery for indeterminate thyroid nodules are influenced by several factors (225), including the estimated presurgical likelihood of malignancy based upon clinical risk factors (>4 cm, family history, and/or history of radiation) (226–229), sonographic pattern (Table 6, Fig. 2) (202–203), cytologic category (Table 8), and ancillary test findings (see molecular testing section [A15–19]). These risk factors, as well as patient preference, presence of contralateral nodularity or coexistent hyperthyroidism, and medical comorbidities, impact decisions regarding thyroid lobectomy with the possible need for subsequent completion thyroideectomy versus total thyroideectomy up front.

Nodules that are cytologically classified as AUS/FLUS or FN and benign using the 167 GEC, or AUS/FLUS and negative using the seven-gene mutation panel have an estimated 5%–6% risk of malignancy (162,163). Nodules that are cytologically classified as FN and negative using the seven-gene mutation panel have an estimated 14% risk of malignancy, which is slightly lower than the risk based upon the Bethesda classification alone (162).

Nodules that are cytologically classified as AUS/FLUS or FN and as suspicious using the 167 GEC have an estimated 37%–44% risk of malignancy, which is slightly higher than the risk based upon the Bethesda classification alone (Table 8) (163,171).

Nodules that are cytologically classified as SUSP cytology and benign using the 167 GEC or negative using the mutation seven-gene panel, also have an estimated 15%–28% risk of malignancy.

In contrast, nodules that are cytologically classified as AUS/FLUS or FN and that are positive for known RAS mutations associated with thyroid carcinoma have an estimated 84% risk of malignancy and should be considered in a similar risk category to cytologically suspicious for malignancy (Table 8) (103,230). Nodules that are cytologically classified as AUS/FLUS or FN or SUSP and that are positive for known BRAFV600E, RET/PTC, or PAX8/PPARγ mutations have an estimated risk of malignancy of >95% and should be considered in a similar category to cytologically diagnosed thyroid carcinoma (198,210,211).

Sonographic pattern of nodules with AUS/FLUS cytology may aid in risk stratification and management. In one study, sonographically benign or seemingly very low risk nodules with AUS/FLUS cytology were noted to be malignant in only 8% of cases, compared to 58% when sonographic suspicious was low or intermediate, and 100% when sonographic suspicion of malignancy was high (203). Another study supported this finding with sonographic findings highly suspicious for malignancy (taller than wide shape, irregular borders, and/or marked hypoechoogenicity) having >90% specificity and PPV for malignancy (202).

The risks of total thyroideectomy are significantly greater than that for thyroid lobectomy, with a recent meta-analysis suggesting a pooled relative risk (RR) significantly greater for all complications, including recurrent laryngeal nerve injury (transient RR = 1.7, permanent RR = 1.9), hypocalcemia (transient RR = 10.7, permanent RR = 3.2), and hemorrhage/hematoma (RR = 2.6) (231). Further, total thyroideectomy is associated with the rare but potential risk of bilateral recurrent laryngeal nerve injury necessitating tracheostomy. Surgeon experience likely influences the risks of thyroideectomy, with higher volume surgeons having lower complication rates (232,233).

Total thyroideectomy necessitates thyroid hormone replacement, while lobectomy is associated with an incidence of postoperative biochemical hypothyroidism estimated at 22%, with clinical or overt hypothyroidism estimated at 4% (234). A significantly increased risk of hypothyroidism following lobectomy has been reported in the presence of autoimmune thyroid disease (e.g., as reflected by the presence of thyroid antibodies) or high normal/elevated TSH (231,234). Hypothyroidism is not an indication for thyroideectomy, and its use as justification for total thyroideectomy over lobectomy should be weighed against the higher risks associated with total thyroideectomy. In contrast, coexistent hyperthyroidism may be an indication for total thyroideectomy depending upon the etiology.

Thyroid lobectomy (hemithyroidectomy) provides definitive histological diagnosis and complete tumor removal for cytologically indeterminate nodules with a lower risk of complications compared to total thyroideectomy and may be sufficient for smaller, solitary intrathyroidal nodules that ultimately prove malignant. As the likelihood of malignancy increases, the potential need for a second operation also increases, if the cytologically indeterminate nodule ultimately proves malignant and if completion thyroideectomy would be recommended. Intraoperative evaluation, with or without frozen section, can occasionally confirm malignancy at the time of lobectomy allowing for conversion to total thyroideectomy if indicated. Frozen section is most helpful if the histopathologic diagnosis is classic PTC, whereas its impact is low in follicular variant of PTC and FTC. The individual patient must weigh the relative advantages and disadvantages of thyroid lobectomy with possible total thyroideectomy or...
RECOMMENDATION 21

(A) Patients with multiple thyroid nodules $\geq 1$ cm should be evaluated in the same fashion as patients with a solitary nodule $\geq 1$ cm, excepting that each nodule that is $>1$ cm carries an independent risk of malignancy and therefore multiple nodules may require FNA.

(Strong recommendation, Moderate-quality evidence)

(B) When multiple nodules $\geq 1$ cm are present, FNA should be performed preferentially based upon nodule sonographic pattern and respective size cutoff (Table 6, Fig. 2).

(Strong recommendation, Moderate-quality evidence)

(C) If none of the nodules has a high or moderate suspicion sonographic pattern, and multiple sonographically similar very low or low suspicion pattern nodules coalesce with no intervening normal parenchyma, the likelihood of malignancy is low and it is reasonable to aspirate the largest nodules ($\geq 2$ cm) or continue surveillance without FNA.

(Weak recommendation, Low-quality evidence)

RECOMMENDATION 22

A low or low-normal serum TSH concentration in patients with multiple nodules may suggest that some nodule(s) may be autonomous. In such cases, a radionuclide (preferably $^{123}$I) thyroid scan should be considered and directly compared to the US images to determine functionality of each nodule $\geq 1$ cm. FNA should then be considered only for those isofunctioning or nonfunctioning nodules, among which those with high suspicion sonographic pattern should be aspirated preferentially.

(Weak recommendation, Low-quality evidence)

Patients with multiple thyroid nodules have the same risk of malignancy as those with solitary nodules (32,74). However, when evaluating the risk of cancer per individual nodule, one large study found that a solitary nodule had a higher likelihood of malignancy than did a nonsolitary nodule ($p < 0.01$), although in agreement with the other studies the risk of malignancy per patient was the same and independent of the number of nodules (77). A recent systematic review and meta-analysis confirmed the slightly higher risk of malignancy in a solitary nodule compared with an individual nodule in a MNG. However, this appeared to hold true mostly outside of the United States and in iodine-deficient populations (235). A diagnostic US should be performed to evaluate the sonographic risk pattern of each nodule, but if only the “dominant” or largest nodule is aspirated, the thyroid cancer may be missed (74). Therefore, multiple thyroid nodules $\geq 1$ cm may require aspiration, based on sonographic pattern (Recommendation 8, Table 6, Fig. 2) to fully exclude clinically relevant thyroid cancer. Radionuclide scanning may also be considered in patients with multiple thyroid nodules with the goal of identifying and aspirating appropriate hypofunctioning nodules. Such imaging may prove especially useful when the serum TSH is below or in the low-normal range. Similarly, sonographic risk assessment of each nodule can assist in identifying those nodules with the highest likelihood of cancer (see section [A10]).
nodule size, which is equivalent to a 20% increase in two of the three nodule dimensions. If a 50% volume increase cutoff is applied, only 4%–10% of nodules were reported to be larger at a mean of 18 months (133,238). However, using cutoffs of a 15% volume increase based upon internally assessed interobserver coefficients of variation, published series report that 32%–50% of nodules increase in size over a 4–5 year period (239,240). Because of the stringent methodology of these studies, adoption of a 15% volume increase as statistically significant is not practically applicable.

A recent 5-year prospective multicenter study evaluated outcomes of 1597 nodules from 992 patients with either cytologically or sonographically benign nodules (241). Nodules 1 cm or larger underwent US-guided FNA and subcentimeter nodules were defined as sonographically benign based upon imaging characteristics equivalent to the ATA low or very low suspicion US patterns. All nodules were followed by annual US exams. The false-negative rate of a benign cytology diagnosis was 1.1%. Of the four missed cancers, on baseline US imaging three were hypoechoic and solid and one was isoechoic with microcalcifications; none was spongiform or mixed cystic solid and noncalcified (ATA very low suspicion pattern). During sonographic surveillance, repeat FNA was prompted by either growth (two nodules) or development of a new suspicious sonographic feature (two nodules). In addition, the shortest time interval to detect change and repeat the FNA was 2 years. Another critical observation from this study was that only one cancer was detected in 5 years among the 852 subcentimeter nodules classified as sonographically benign at baseline. This cancer was identified on the 5-year follow-up US, when its composition changed from mixed cystic/solid to hypoechoic solid with irregular margins prompting FNA (241). Currently, there are no follow-up studies of nodule growth that extend observation beyond 5 years to help inform decision-making about long-term surveillance. Additional research would be valuable because indefinite follow-up of nodules with benign cytology is costly and may be unnecessary.

Recent investigations of repeat US-guided FNA in nodules with initial benign cytology show higher detection rates for missed malignancy for those nodules with a high suspicion sonographic pattern rather than size increase (236,242). Kwak et al. (236) reported a significantly higher malignancy rate of 20.4% in nodules with benign cytology that exhibited either marked hypoechoogenicity, irregular borders, microcalcifications, or a taller nodules with benign cytology diagnoses should be determined by risk stratification based upon US pattern as defined in Recommendation 8. If follow-up US for surveillance is performed and the nodule size is stable, the utility of subsequent US imaging for detection of potential malignancy by nodule growth assessment is very low and if performed, the time interval for any additional US exam should be at least as long that between the initial benign FNA cytology result and first follow-up. However, even if a repeat US is not indicated based on a benign cytology, US pattern, or stability in nodule size, larger nodules may require monitoring for growth that could result in symptoms and thus prompt surgical intervention despite benign cytology.

One recent study evaluated the long-term consequences of a false-negative benign cytology (140). A total of 1369 patients with 2010 cytologically benign thyroid nodules were followed for a mean of 8.5 years. Eighteen false-negative cases were identified, although only a subset of patients underwent repeat FNA or thyroid surgery. Thirty deaths were documented in the entire cohort over this time period and none were attributable to thyroid cancer. These data support that an initial benign cytology conveys an overall excellent prognosis and a conservative follow-up strategy is reasonable.

**[A26] Follow-up for nodules that do not meet FNA criteria**

### RECOMMENDATION 24

Nodules may be detected on US that do not meet criteria for FNA at initial imaging (Recommendation 8). The strategy for sonographic follow-up of these nodules should be based upon the nodule’s sonographic pattern.

(A) Nodules with high suspicion US pattern: repeat US in 6–12 months.

(Weak recommendation, Low-quality evidence)

(B) Nodules with low to intermediate suspicion US pattern: consider repeat US at 12–24 months.

(Weak recommendation, Low-quality evidence)

(C) Nodules >1 cm with very low suspicion US pattern (including spongiform nodules) and pure cyst: the utility and time interval of surveillance US for risk of malignancy is not known. If US is repeated, it should be at ≥24 months.

(No recommendation, Insufficient evidence)

(D) Nodules ≤1 cm with very low suspicion US pattern (including spongiform nodules) and pure cysts do not require routine sonographic follow-up.

(Weak recommendation, Low-quality evidence)

Ultrasound studies demonstrate that up to 50% of adults have thyroid nodules. The vast majority of these are subcentimeter, and FNA evaluation is generally not indicated. In addition, based upon both sonographic pattern and size cutoffs (Recommendation 8), many nodules >1 cm may also be followed without FNA. Although no prospective studies address the optimal cost-effective surveillance strategy for these nodules that have not undergone FNA, a recent study by Durante et al. (241) confirms that subcentimeter thyroid nodules corresponding to the ATA very low suspicion risk pattern are highly

...
unlikely to change during 5-year sonographic follow-up, and the risk of malignancy is exceedingly low. The findings from studies correlating sonographic features and malignancy risk in aspirated nodules can be extrapolated to inform a follow-up strategy for this group of nodules that do not meet FNA criteria at the time of their initial detection. For example, the interval for follow-up sonography for a nodule that is hypoechoic and taller than wide should be shorter than that for an isoechoic solid nodule with smooth borders. US-guided FNA should be performed based upon follow-up US imaging if the nodule subsequently meets criteria based upon Recommendation 8.

[A27] What is the role of medical or surgical therapy for benign thyroid nodules?

RECOMMENDATION 25

Routine TSH suppression therapy for benign thyroid nodules in iodine sufficient populations is not recommended. Though modest responses to therapy can be detected, the potential harm outweighs benefit for most patients.

(Strong recommendation, High-quality evidence)

RECOMMENDATION 26

Individual patients with benign, solid, or mostly solid nodules should have adequate iodine intake. If inadequate dietary intake is found or suspected, a daily supplement (containing 150 μg iodine) is recommended.

(Strong recommendation, Moderate-quality evidence)

RECOMMENDATION 27

(A) Surgery may be considered for growing nodules that are benign after repeat FNA if they are large (>4 cm), causing compressive or structural symptoms, or based upon clinical concern.

(Weak recommendation, Low-quality evidence)

(B) Patients with growing nodules that are benign after FNA should be regularly monitored. Most asymptomatic nodules demonstrating modest growth should be followed without intervention.

(Strong recommendation, Low-quality evidence)

RECOMMENDATION 28

Recurrent cystic thyroid nodules with benign cytology should be considered for surgical removal or percutaneous ethanol injection (PEI) based on compressive symptoms and cosmetic concerns. Asymptomatic cystic nodules may be followed conservatively.

(Weak recommendation, Low-quality evidence)

RECOMMENDATION 29

There are no data to guide recommendations on the use of thyroid hormone therapy in patients with growing nodules that are benign on cytology.

(No recommendation, Insufficient evidence)

Evidence from multiple prospective, RCTs, and from three meta-analyses suggest that thyroid hormone supplementation in doses that suppress the serum TSH to subnormal levels may result in a decrease in nodule size and may prevent the appearance of new nodules in regions of the world with borderline low iodine intake (239,243–245). However, the effect is modest, with most studies suggesting an average 5%–15% reduction in nodule volume when treated with suppressive levothyroxine (LT4) therapy for 6–18 months. Two high-quality meta-analyses confirm that six to eight patients will require suppressive LT4 therapy to achieve one successful treatment response (246,247). The extent of TSH suppression achieved in high-quality studies is variable, though the majority suppressed TSH to <0.2 mIU/L, with many to <0.1 mIU/L. Hyperthyroidism to this degree has been significantly associated with an increased risk of cardiac arrhythmias and osteoporosis, as well as adverse symptomatology. Together, these data confirm that LT4 suppressive therapy demonstrates modest (though usually clinically insignificant) efficacy in nodule volume reduction, but increases the risk of adverse consequences related to iatrogenic thyrotoxicosis. One large prospective, randomized trial demonstrated that sufficient dietary iodine intake (150 μg daily) also reduced nodule size slightly more than placebo (248). The consumption of adequate dietary iodine is recommended for all adults and is without harm when not excessive. Data supporting LT4 therapy in non-TSH-suppressive doses for prevention of thyroid nodule growth are incomplete. One recent cohort analysis suggested non-suppressive doses of LT4 therapy conferred protection from nodule growth over time (249). However, the nonblinded, non-randomized nature of the trial precludes broad translation of the data, and the efficacy of nonsuppressive LT4 remains unproven.

Cystic nodules that are cytologically benign can be monitored for recurrence (fluid reaccumulation), which can be seen in 60%–90% of patients (250,251). For those patients with subsequent recurrent symptomatic cystic fluid accumulation, surgical removal, generally by hemithyroidectomy, or PEI are both reasonable strategies. Four controlled studies demonstrated a 75%–85% success rate after PEI compared with a 7%–38% success rate in controls treated by simple cyst evacuation or saline injection. Success was achieved after an average of two PEI treatments. Complications included mild to moderate local pain, flushing, dizziness, and dysphonia (250–253). Surgery may be considered for growing solid nodules that are benign on repeat cytology if they are large (>4 cm), causing compressive or structural symptoms, or based upon clinical concern (254,255).

A28 How should thyroid nodules in pregnant women be managed?

A29 FNA for thyroid nodules discovered during pregnancy

RECOMMENDATION 30

(A) FNA of clinically relevant thyroid nodules (refer to section [A10]) should be performed in euthyroid and hypothyroid pregnant women.

(Strong recommendation, Moderate-quality evidence)

(B) For women with suppressed serum TSH levels that persist beyond 16 weeks gestation, FNA may be deferred until after pregnancy and cessation of lactation. At that
time, a radionuclide scan can be performed to evaluate nodule function if the serum TSH remains suppressed.

(Strong recommendation, Moderate-quality evidence)

It is uncertain if thyroid nodules discovered in pregnant women are more likely to be malignant than those found in nonpregnant women, since there are no population-based studies to address this question. Pregnancy does not appear to modify microscopic cellular appearance, thus standard diagnostic criteria should be applied for cytologic evaluation (256). Serial evaluation of nodules throughout pregnancy has demonstrated that thyroid nodules will enlarge slightly throughout gestation, though this does not imply malignant transformation (257). The recommended evaluation of a clinically relevant nodule in a pregnant patient is thus the same as for a non-pregnant patient, with the exception that a radionuclide scan is contraindicated. In addition, for patients with nodules diagnosed as DTC by FNA during pregnancy, delaying surgery until after delivery does not affect outcome (258). Surgery performed during pregnancy is associated with greater risk of complications, longer hospital stays, and higher costs (259).

[A30] Approaches to pregnant patients with malignant or indeterminate cytology

RECOMMENDATION 31

PTC discovered by cytology in early pregnancy should be monitored sonographically. If it grows substantially (as defined in section [A24]) before 24–26 weeks gestation, or if US reveals cervical lymph nodes that are suspicious for metastatic disease, surgery should be considered during pregnancy. However, if the disease remains stable by midgestation, or if it is diagnosed in the second half of pregnancy, surgery may be deferred until after delivery.

(Weak recommendation, Low-quality evidence)

If FNA cytology is consistent with PTC, surgery is generally recommended. However, the decision to perform such surgery either during pregnancy or after delivery must be individualized. If surgery is not performed, the utility of thyroid hormone therapy targeted to lower serum TSH levels to improve the prognosis of thyroid cancer diagnosed during gestation is not known. Because higher serum TSH levels may be correlated with a more advanced stage of cancer at surgery (260), if the patient’s serum TSH is >2 mU/L, it may be reasonable to initiate thyroid hormone therapy to maintain the TSH between 0.3 to 2.0 mU/L for the remainder of gestation.

Most data confirm that the prognosis of women with well-differentiated thyroid cancer identified but not treated during pregnancy is similar to that of nonpregnant patients. Because of this, surgery in most pregnant patients is deferred until postpartum (258,261), and no further testing is required. However, some studies differ from these findings. Two Italian cohort studies have investigated women diagnosed with DTC in relation to the timing of pregnancy. Messuti et al. (262) noted a statistically higher rate of persistence/recurrence when DTC was diagnosed during pregnancy or within 2 years postpartum. However, the stimulated Tg was found to be >10 ng/mL during 131I ablation in many cases, suggesting the extent of thyroidectomy and/or tumor resection may have been limited in this cohort and therefore contributed to biochemical persistence of disease. Vannucchi et al. (263) followed a small cohort of 10 patients with DTC during pregnancy or within 1 year post partum, again noting a large rate of persistent disease (60%) compared to nonpregnant controls (4.2%–13.1%). Similarly, the majority of cases with persistent disease were attributable to biochemical elevations in Tg or anti-Tg antibodies, again raising the question of whether the extent of initial resection was limited in comparison to nonpregnant controls. Given the likelihood that biochemical persistence could be attributable to an increased size of remnant tissue or incomplete surgical resection in both studies, these data should not refute previous, larger analyses showing no increased recurrence rates when DTC is diagnosed during pregnancy.

Theoretically, molecular marker analysis could be helpful in the evaluation of DTC or clinically relevant, cytologically indeterminate thyroid nodules detected during pregnancy. However, the application of molecular testing in pregnant women with indeterminate cytology remains uncertain. There are no published data validating the performance of any molecular marker in this population. Therefore, the committee cannot recommend for or against their use in pregnant women. However, it is theoretically possible that changes in a nodule’s RNA expression may occur during gestation altering performance of the 167 GEC while the seven-gene mutational panel (BRAF, RAS, PAX8/PPARY, RET/PTC) would be more likely to demonstrate similar performance to that of a nonpregnant population.

When surgery is advised during pregnancy, it is most often because of high-risk clinical or sonographic findings, nodule growth, or change over short duration follow-up or it is based upon physician judgement. To minimize the risk of miscarriage, surgery during pregnancy should be done in the second trimester before 24 weeks gestation (264). However, PTC discovered during pregnancy does not behave more aggressively than that diagnosed in a similar-aged group of nonpregnant women (258,265). A retrospective study of pregnant women with DTC found no difference in either recurrence or survival rates between women operated during or after pregnancy (258). Further, retrospective data suggest that treatment delays of <1 year from the time of thyroid cancer discovery do not adversely affect patient outcome (266). A separate study reported a higher rate of complications in pregnant women undergoing thyroid surgery compared with nonpregnant women (267). If FNA cytology is indeterminate, monitoring may be considered with further evaluation may be delayed until after delivery. Some experts recommend thyroid hormone suppression therapy for pregnant women with FNA suspicious for or diagnostic of PTC, if surgery is deferred until the postpartum period (259).

[81] DIFFERENTIATED THYROID CANCER: INITIAL MANAGEMENT GUIDELINES

Differentiated thyroid cancer, arising from thyroid follicular epithelial cells, accounts for the vast majority of thyroid cancers. Of the differentiated cancers, papillary cancer comprises about 85% of cases compared to about 12% that have follicular histology, including conventional and oncocytic (Hürthle cell) carcinomas, and <3% that are poorly
differentiated tumors (268). In general, stage for stage, the prognoses of PTC and follicular cancer are similar (266,269).

[B2] Goals of initial therapy of DTC

The basic goals of initial therapy for patients with DTC are to improve overall and disease-specific survival, reduce the risk of persistent/recurrent disease and associated morbidity, and permit accurate disease staging and risk stratification, while minimizing treatment-related morbidity and unnecessary therapy. The specific goals of initial therapy are to

1. Remove the primary tumor, disease that has extended beyond the thyroid capsule, and clinically significant lymph node metastases. Completeness of surgical resection is an important determinant of outcome, while residual metastatic lymph nodes represent the most common site of disease persistence/recurrence (270–272).

2. Minimize the risk of disease recurrence and metastatic spread. Adequate surgery is the most important treatment variable influencing prognosis, while RAI treatment, TSH suppression, and other treatments each play adjunctive roles in at least some patients (273–275).

3. Facilitate postoperative treatment with RAI, where appropriate. For patients undergoing RAI remnant ablation, or RAI treatment of presumed (adjuvant therapy) or known (therapy) residual or metastatic disease, removal of all normal thyroid tissue is an important element of initial surgery (276).

4. Permit accurate staging and risk stratification of the disease. Because disease staging and risk stratification should be used to guide initial prognostication, disease management, and follow-up strategies, accurate postoperative risk assessment is a crucial element in the management of patients with DTC (277,278).

5. Permit accurate long-term surveillance for disease recurrence.

6. Minimize treatment-related morbidity. The extent of surgery and the experience of the surgeon both play important roles in determining the risk of surgical complications (232,233,279,280).

[B3] What is the role of preoperative staging with diagnostic imaging and laboratory tests?

[B4] Neck imaging—ultrasound

- RECOMMENDATION 32

(A) Preoperative neck US for cervical (central and especially lateral neck compartments) lymph nodes is recommended for all patients undergoing thyroidectomy for malignant or suspicious for malignancy cytologic or molecular findings.

(Strong recommendation, Moderate-quality evidence)

(B) US-guided FNA of sonographically suspicious lymph nodes 28–10 mm in the smallest diameter should be performed to confirm malignancy if this would change management.

(Strong recommendation, Moderate-quality evidence)

(C) The addition of FNA-Tg washout in the evaluation of suspicious cervical lymph nodes is appropriate in select patients, but interpretation may be difficult in patients with an intact thyroid gland.

(Weak recommendation, Low-quality evidence)

Differentiated thyroid carcinoma (particularly papillary carcinoma) involves cervical lymph node metastases in 20%–50% of patients in most series using standard pathologic techniques (84,145,281–283), and it may be present even when the primary tumor is small and intrathyroidal (284). The frequency of micrometastases (<2 mm) may approach 90%, depending on the sensitivity of the detection method (285,286). However, the clinical implications of micrometastases are likely less significant compared to macrometastases. Preoperative US identifies suspicious cervical adenopathy in 20%–31% of cases, potentially altering the surgical approach (287,288) in as many as 20% of patients (289–291). However, preoperative US identifies only half of the lymph nodes found at surgery, due to the presence of the overlying thyroid gland (292).

Sonographic features suggestive of abnormal metastatic lymph nodes include enlargement, loss of the fatty hilum, a rounded rather than oval shape, hyperechogenicity, cystic change, calcifications, and peripheral vascularity (Table 7). No single sonographic feature is adequately sensitive for detection of lymph nodes with metastatic thyroid cancer. One study correlated the sonographic features acquired 4 days preoperatively directly with the histology of 56 cervical lymph nodes identified in 19 patients. Some of the most specific criteria were short axis >5 mm (96%), presence of cystic areas (100%), presence of hyperechogenic punctuations representing either colloid or microcalcifications (100%), and peripheral vascularity (82%). Of these, the only one with sufficient sensitivity was peripheral vascularity (86%). The others had sensitivities of <60% and would not be adequate to use as a single criterion for identification of malignant involvement (292). As shown by earlier studies (293,294), the ultrasonographic feature with the highest sensitivity is absence of a hilum (100%), but this has a low specificity of 29%. Microcalcifications have the highest specificity; any lymph nodes with microcalcifications should be considered abnormal (292) (Table 7).

The location of the lymph nodes may also be useful for decision-making. Malignant lymph nodes are much more likely to occur in levels III, IV, and VI than in level II (Fig. 3) (292,294), although this may not be true for PTC tumors arising in the upper pole of the thyroid, which have a higher propensity to demonstrate skip metastases to levels III and II (295). Figure 3 illustrates the delineation of cervical lymph node levels I through VI.

Confirmation of malignancy in lymph nodes with a suspicious sonographic appearance is achieved by US-guided FNA aspiration for cytology and/or measurement of Tg in the needle washout. A Tg concentration <1 ng/mL is reassuring, and the probability of N1 disease increases with higher Tg levels (296). This FNA measurement of Tg is likely valid even in patients with circulating anti-Tg autoantibodies (297,298), although one study challenges the validity of this measurement in patients with anti-Tg autoantibodies (299). Tg washout may be helpful, particularly in cases in which the lymph nodes are cystic, cytologic evaluation of the lymph node is inadequate, or the cytologic and sonographic evaluations are divergent (i.e., normal cytologic biopsy of
a large lymph node with microcalcifications) (300). In a retrospective study of 241 lymph nodes in 220 patients who underwent US-guided FNA with Tg in FNA (FNA-Tg) washout fluid measurements for suspicious lymph nodes, additional FNA-Tg helped to diagnose a metastatic lymph node with one or two suspicious US features but did not offer incremental benefit for those lymph nodes with highly suspicious US features in which FNA alone was sufficient for diagnosis. Two recent systematic reviews showed that false-positive Tg washout may occur, particularly in lymph nodes in the central compartment when the thyroid gland is still present (301,302). The review by Pak et al. (302) suggests that an FNA-Tg cutoff of 32 ng/mL has the best sensitivity and specificity in patients with an intact thyroid gland. Others have suggested interpreting the FNA-Tg in context of the serum Tg and TSH in these patients (303,304). There is no standardization of FNA-Tg procedures or assays to date, which makes this additional diagnostic tool sometimes difficult to interpret (305). Future standardization including matrix type (phosphate-buffered saline, Tg-free serum, etc.) and volume of diluent matrix would help with interpretation of a Tg washout.

Accurate staging is important in determining the prognosis and tailoring treatment for patients with DTC. However, unlike many tumor types, the presence of metastatic disease does not obviate the need for surgical excision of the primary tumor in DTC (306). Because metastatic disease may respond to RAI therapy, removal of the thyroid as well as the primary tumor and accessible loco-regional disease remains an important component of initial treatment even in most patients with metastatic disease.

**RECOMMENDATION 33**

(A) Preoperative use of cross-sectional imaging studies (CT, MRI) with intravenous (IV) contrast is recommended as an adjunct to US for patients with clinical suspicion for advanced disease, including invasive primary tumor, or clinically apparent multiple or bulky lymph node involvement.

(Strong recommendation, Low-quality evidence)

(B) Routine preoperative 18F-FDG-PET scanning is not recommended.

(Strong recommendation, Low-quality evidence)

Since US evaluation is operator dependent and cannot always adequately image deep anatomic structures and those acoustically shadowed by bone or air, alternative imaging procedures may be preferable or useful as an adjunct in
some clinical settings. Patients displaying bulky or widely distributed nodal disease on initial US examination may present with involvement of nodal regions beyond typical cervical regions, some of which maybe difficult to visualize on routine preoperative US, including the mediastinum, infraclavicular, retropharyngeal, and parapharyngeal regions. In a study of 37 consecutive patients who had preoperative CT and US and subsequently underwent total thyroidectomy and neck dissection, the sensitivity of CT was better than US for the evaluation central and lateral compartment lymph nodes examined together (77% vs 62%, p = 0.002), but there were no differences between the two imaging modalities when the central and lateral compartments were examined separately (307). In a series of 299 consecutively registered patients with pathologically proven PTC who underwent preoperative CT and US, US was more accurate than CT in predicting extrathyroidal tumor extension and multifocal bilobar disease (p < 0.05). The accuracy of staging was better overall with US (p < 0.01), and US had greater sensitivity than CT at predicting lateral compartment metastases (p = 0.041) (308).

However, another study showed that combined preoperative mapping with US and CT was superior to US alone in the preoperative detection of nodal disease, especially in the central neck (309). The sensitivities of MRI and PET for the detection of cervical lymph node metastases are relatively low (30%–40%) (310). PET can also detect inflammatory lymph nodes, which reduces the specificity of this test in many patients with DTC. MRI can be used in the detection of cervical nodal metastasis. MRI is affected by respiratory artifacts and may be more difficult to interpret than CT scanning by surgeons in the operating room for low-volume nodal disease (311).

Invasive DTC has been reported to occur in 10%–15% of patients at the time of diagnosis (312). For this group of patients, cross-sectional imaging can also be a useful supplement for preoperative planning to accurately delineate the extent of laryngeal, tracheal, esophageal, or vascular involvement (309,313). Endoscopy of the trachea and or esophagus, with or without ultrasonography, at the beginning of the initial operation looking for evidence of intraluminal extension can also be helpful in cases of suspected aerodigestive tract invasion.

Locally invasive primary tumors may be associated with characteristic signs and symptoms including progressive dysphagia, respiratory compromise, hemoptysis, rapid tumor enlargement, significant voice change or the finding of vocal cord paralysis, and mass fixation to the airway or neck structures. Certain sonographic features of the primary tumor, including extrathyroidal extension especially with posterior capsular extension and extension into the mediastinum, may also prompt axial imaging (307). Chest CT is useful in defining the inferior border of disease and in determining the extent to which mediastinal structures are involved in cases with significant caudal spread. CT findings may influence management by indicating the need for sternotomy and/or tracheal or laryngeal resection/reconstruction, which would likely require assembling additional resources and personnel in preparation for surgery. Neck CT with contrast can therefore be useful in delineating the extent of laryngeal, tracheal, and/or esophageal involvement in tumors displaying aggressive local invasion, as well as delineating bulky nodal disease, which may harbor significant extranodal extension that involves muscle and/or blood vessels. Preoperative knowledge of these features of the primary tumor or metastases could significantly influence the surgical plan (314). 18FDG-PET scanning may be sensitive in some patients for neck or mediastinal involvement and may reveal distant metastases as well.

When cross-sectional imaging is performed, use of IV contrast is an important adjunct because it helps to delineate the anatomic relationship between the primary tumor or metastatic disease and these other structures. Iodine is generally cleared within 4–8 weeks in most patients, so concern about iodine burden from IV contrast causing a clinically significant delay in subsequent whole-body scans (WBSs) or RAI treatment after the imaging followed by surgery is generally unfounded (315). The benefit gained from improved anatomic imaging generally outweighs any potential risk of a several week delay in RAI imaging or therapy. When there is concern, a urinary iodine to creatinine ratio can be measured.

RECOMMENDATION 34

Routine preoperative measurement of serum Tg or anti-Tg antibodies is not recommended.

(Weak recommendation, Low-quality evidence)

Data from a systematic review and meta-analysis suggested that high preoperative concentrations of serum Tg may predict a higher sensitivity for postoperative surveillance with serum Tg (316). Preoperative anti-Tg antibodies do not appear to be an independent preoperative predictor of stage in patients with DTC, but the evidence is limited. In a cross-sectional analysis of 1770 patients with perioperative anti-Tg antibodies status data in the National Thyroid Cancer Treatment Cooperative Study (a large thyroid cancer registry that included 11 North American centers and enrolled patients between 1987 and 2011), serum anti-Tg antibody status was not significantly associated with stage of disease on multivariate analysis, or with disease-free or overall survival on univariate or multivariate analyses (317). Evidence that preoperative measurement of serum Tg impacts patient management or outcomes is not yet available.

RECOMMENDATION 35

(A) For patients with thyroid cancer >4 cm, or with gross extrathyroidal extension (clinical T4), or clinically apparent metastatic disease to nodes (clinical N1) or distant sites (clinical M1), the initial surgical procedure should include a near-total or total thyroidectomy and gross removal of all primary tumor unless there are contraindications to this procedure.

(Strong recommendation, Moderate-quality evidence)

(B) For patients with thyroid cancer >1 cm and <4 cm without extrathyroidal extension, and without clinical evidence of any lymph node metastases (cN0), the initial surgical procedure can be either a bilateral procedure (near-total or total thyroidectomy) or a unilateral procedure.
Thyroid lobectomy alone may be sufficient initial treatment for low-risk papillary and follicular carcinomas; however, the treatment team may choose total thyroidectomy to enable RAI therapy or to enhance follow-up based upon disease features and/or patient preferences.

(Strong recommendation, Moderate-quality evidence)

(C) If surgery is chosen for patients with thyroid cancer <1 cm without extrathyroidal extension and cN0, the initial surgical procedure should be a thyroid lobectomy unless there are clear indications to remove the contralateral lobe. Thyroid lobectomy alone is sufficient treatment for small, unifocal, intrathyroidal carcinomas in the absence of prior head and neck radiation, familial thyroid carcinoma, or clinically detectable cervical nodal metastases.

(Strong recommendation, Moderate-quality evidence)

Surgery for thyroid cancer is an important element of a multifaceted treatment approach. The operation must be compatible with the overall treatment strategy and follow-up plan recommended by the managing team. Consideration should be given to referring patients with high-risk features (clinical N1 disease, concern for recurrent laryngeal nerve [RLN] involvement, or grossly invasive disease) to experienced surgeons, as both completeness of surgery and experience of the surgeon can have a significant impact on clinical outcomes and complication rates (232,233,279,280). Previous guidelines have endorsed total thyroidectomy as the primary initial surgical treatment option for nearly all DTCs >1 cm with or without evidence of loco-regional or distant metastases (25). This was based on retrospective data suggesting that a bilateral surgical procedure would improve survival (318), decrease recurrence rates (319–321), allow for routine use of RAI remnant ablation, and facilitate detection of recurrent/persistent disease during follow-up. However, recent data have demonstrated that in properly selected patients, clinical outcomes are very similar following unilateral or bilateral thyroid surgery (322–326). Furthermore, since the requirement for routine use of RAI ablation was one of the major reasons given in support of total thyroidectomy in low to intermediate risk patients, our current more selective approach to RAI ablation in these patients requires a critical reassessment of this indication. In some patients, the presence of the remaining lobe of the gland may obviate the lifelong need for exogenous thyroid hormone therapy. Finally, as our follow-up management paradigm has moved away from diagnostic whole body RAI scanning and toward a greater reliance on neck ultrasonography and serial serum Tg measurements (even in patients that did not receive RAI remnant ablation), we must also question whether total thyroidectomy and RAI remnant ablation is required to facilitate follow-up in low to intermediate risk patients.

In an analysis of 52,173 PTC patients diagnosed between 1985 and 1998 from the National Cancer Data Base (43,227 receiving total thyroidectomy, 8946 undergoing lobectomy), Bilimoria et al. (318) demonstrated a slightly higher 10-year relative overall survival for total thyroidectomy as opposed to thyroid lobectomy (98.4% vs. 97.1%, respectively, p < 0.05) and a slightly lower 10-year recurrence rate (7.7% vs. 9.8%, respectively, p < 0.05). When analyzed by size of the primary tumor, statistically significant differences in survival and recurrence were seen for all sizes >1 cm based on the extent of initial surgery. However, data on extrathyroidal extension, completeness of resection, and other comorbid conditions, which could have had a major impact on survival and recurrence risk, were not available. Therefore, it is unclear how often lobectomy was done based on proper selection of low to intermediate risk patients versus how often lobectomy was done in high-risk patients because of comorbid conditions, inability to obtain a complete resection, or status of the contralateral RLN. This is an important distinction because thyroid lobectomy patients were found to have extrathyroidal extension in 7% of cases (325), underwent external beam radiation therapy (EBRT) in 1%–2% (324), and RAI therapy in 12%–18% (318,325), and high-risk features were present in 8% (325). Given the small magnitude of differences reported for survival and recurrence between the total thyroidectomy and the lobectomy patients, it is quite possible that the slightly poorer outcomes seen in the lobectomy group could have been influenced by lobectomy patients with concurrent high risk features. Adam et al. (327) performed an updated analysis of 61,775 patients in the National Cancer Database who underwent thyroid surgery between 1998 and 2006. The researchers demonstrated that the overall survival advantage seen for patients with 1–4 cm PTC who underwent thyroidectomy in the study by Bilimoria et al. (318) disappeared when further adjustment was made for additional variables related to complexity and severity of illness. This lack of overall survival advantage was also seen when the group was subdivided into patients with 1–2 cm and 2–4 cm PTC.

Previously, Haigh et al. (325) had analyzed 5432 PTC patients from the SEER database (4612 receiving total thyroidectomy and 820 undergoing lobectomy) and found no difference in 10-year overall survival between total thyroidectomy and thyroid lobectomy when risk stratified by the AMES classification system. Interestingly, patients selected for thyroid lobectomy included 7% with extrathyroidal extension, 1% with distant metastases, and 5% with primary tumors >5 cm, and 8% were classified as having high risk based on AMES.

More recently, two additional studies have analyzed the SEER database, and both have failed to demonstrate a significant difference in survival when comparing total thyroidectomy with thyroid lobectomy (323,324). Barney et al. (323) included 23,605 DTC patients diagnosed between 1983 and 2002 (12,598 with total thyroidectomy, 3266 with lobectomy) and found no difference in 10-year overall survival (90.4% for total thyroidectomy vs. 90.8% for lobectomy) or 10-year cause-specific survival (96.8% for total thyroidectomy vs. 98.6% for lobectomy). Furthermore, in a multivariate analysis that included age, T, N, M, sex, year of diagnosis, extent of surgery, and RAI use, no difference in overall survival or cause specific survival was seen with respect to the extent of initial surgery. Mendelsohn et al. (324) analyzed 22,724 PTC patients diagnosed between 1998 and 2001 (16,760 with total thyroidectomy, 5964 with lobectomy) and found no differences in overall survival or disease-specific survival in a comparison of total thyroidectomy with lobectomy. Interestingly, of the patients that had lobectomy, 1.6% received external beam radiation therapy, 16% had extrathyroidal extension, 9% of tumors were >4 cm, and 20% received RAI ablation (once again indicating that lobectomy was done in some high-risk patients).
Consistent with the SEER data analyses, two single-center studies also confirmed that lobectomy is associated with excellent survival in properly selected patients (322,326). After a median follow-up of 8 years, only one disease-specific death was seen in a cohort of 889 PTC patients with T1–T2 tumors treated with either total thyroidectomy \((n = 528)\) or lobectomy \((n = 361)\) (326). Furthermore, Matsumu et al. (322) reported a cause-specific survival rate of 98% after a median of 17 years of follow-up in properly select PTC patients treated with lobectomy and ipsilateral neck dissection.

Given the propensity for PTC to be multifocal (often involving both lobes), it is not surprising that some studies have demonstrated a lower risk of loco-regional disease recurrence following total thyroidectomy as compared to thyroid lobectomy (319–321). However, with proper patient selection, loco-regional recurrence rates of less than 1%–4% and completion thyroidectomy rates of <10% can be achieved following thyroid lobectomy (326,328). Furthermore, the few recurrences that develop during long-term follow-up are readily detected and appropriately treated with no impact on survival (322,326,328).

Therefore, we conclude that in properly selected low- to intermediate-risk patients (patients with unifocal tumors <4 cm, and no evidence of extrathyroidal extension or lymph node metastases by examination or imaging), the extent of initial thyroid surgery probably has little impact on disease-specific survival. While recurrence rates can be quite low in these patients, it is likely that the lowest rates of recurrence during long-term follow-up would be associated with a total thyroidectomy. But since salvage therapy is quite effective in the few patients that recur after thyroid lobectomy, a conservative management approach to completion surgery, accepting a slightly higher risk of loco-regional recurrence, is a reasonable management strategy. Finally, a more selective use of RAI coupled with a greater reliance on neck US and serial serum Tg measurements for detection of recurrent disease is likely to significantly decrease the mandate for total thyroidectomies in low- and intermediate-risk patients done solely to facilitate RAI remnant ablation and follow-up.

Near-total or total thyroidectomy is necessary if the overall strategy is to include RAI therapy postoperatively, and thus is recommended if the primary thyroid carcinoma is >4 cm, if there is gross extrathyroidal extension, or if regional or distant metastases are clinically present. For tumors that are between 1 and 4 cm in size, either a bilateral thyroidectomy (total or near-total) or a unilateral procedure (thyroid lobectomy) may be suitable as treatment plan. Older age (>45 years), contralateral thyroid nodules, a personal history of papillary thyroid carcinoma with clinically uninvolved central neck lymph nodes (cN0) who have advanced primary tumors (T3 or T4) or clinically involved lateral neck nodes (cN1b), or if the information will be used to plan further steps in therapy.

The relationship between surgeon volume and patient outcomes has been studied extensively over the last 20 years. Institutional studies examining outcomes following thyroidectomy by high-volume surgeons have been published demonstrating overall safety. In one of the first studies examining the relationship between surgeon volume and thyroidectomy outcomes at a state level, Sosa et al. (232) found a strong association between higher surgeon volume and favorable patient outcomes, especially with regard to RLN injury and wound complications. This was especially pronounced for patients undergoing total thyroidectomy for thyroid cancer. Others have made similar observations (233,329,330).

In a recent study of patients undergoing thyroidectomy in the Health Care Utilization Project Nationwide Inpatient Sample (HCUP-NIS), surgeons were divided into categories of low (<10 cases/year; encompassing 6072 surgeons), intermediate (10–100 cases/year; 11,544 surgeons), and high volume (>100 cases/year; 4009 surgeons) (331). Over 80% of thyroid resections were performed by low- and intermediate-volume surgeons. On average, high-volume surgeons had the lowest complication rates for patients who underwent total thyroidectomy for cancer at 7.5%; intermediate-volume surgeons had a rate of 13.4%, and low-volume surgeons, 18.9% \((p < 0.001)\).

From robust population-level data such as these, it can be concluded that referral of patients to high-volume thyroid surgeons is associated, on average, with superior outcomes. However, such referral is not always possible, given the relative scarcity of high-volume surgeons and their geographic distribution. In addition, there are some data suggesting that other factors, such as surgeon age, should be considered (332). Therefore, conclusions at a population level cannot always be applied to individual surgeons and patient circumstances. It may, however, be reasonable to consider sending patients with more extensive disease and concern for grossly invasive disease to a high-volume surgeon experienced in the management of advanced thyroid cancer.

It is worth noting that even high-volume surgeons have a higher overall postoperative complication rate when performing total thyroidectomy compared with lobectomy (333). Using the HCUP-NIS, these authors found that high-volume thyroid surgeons had a complication rate of 7.6% following thyroid lobectomy but a rate of 14.5% following total thyroidectomy. For low-volume surgeons, the complication rates were 11.8% and 24.1%, respectively. Therefore, patients should carefully weigh the relative benefits and risks of total thyroidectomy versus thyroid lobectomy, even when surgery is performed by high-volume surgeons.

[B8] Lymph node dissection

■ RECOMMENDATION 36

(A) Therapeutic central-compartment (level VI) neck dissection for patients with clinically involved central nodes should accompany total thyroidectomy to provide clearance of disease from the central neck.\(^{[270,278,322,326]}\)

(Strong recommendation, Moderate-quality evidence)

(B) Prophylactic central-compartment neck dissection (ipsilateral or bilateral) should be considered in patients with papillary thyroid carcinoma with clinically uninvolved central neck lymph nodes (cN0) who have advanced primary tumors (T3 or T4) or clinically involved lateral neck nodes (cN1b), or if the information will be used to plan further steps in therapy.\(^{[270,278,322,326]}\)

(Weak recommendation, Low-quality evidence)

(C) Thyroidectomy without prophylactic central neck dissection is appropriate for small (T1 or T2), noninvasive,
RECOMMENDATION 37

Therapeutic lateral neck compartmental lymph node dissection should be performed for patients with biopsy-proven metastatic lateral cervical lymphadenopathy.

(Strong recommendation, Moderate-quality evidence)

Regional lymph node metastases are present at the time of diagnosis in a majority of patients with papillary carcinomas and a lesser proportion of patients with follicular carcinomas (290,334,335). Although PTC lymph node metastases are reported by some to have no clinically important effect on outcome in low risk patients, a study of the SEER database found, among 9904 patients with PTC, that lymph node metastases, age >45 years, distant metastasis, and large tumor size significantly predicted poor overall survival outcome in a multivariate analysis (336). All-cause survival at 14 years was 82% for PTC without lymph node metastases and 79% with nodal metastases (p<0.05). Another SEER registry study concluded that cervical lymph node metastases conferred an independent risk of decreased survival, but only in patients with follicular cancer and patients with papillary cancer over age 45 years (337). However, characteristics of the lymph node metastases can further discriminate the risk of recurrence to the patient, especially in patients with clinically evident metastasis, multiple metastases, larger metastases, and/or extracapsular nodal extension (338,339), compared with those with more limited microscopic nodal disease (335). A recent comprehensive analysis of the National Cancer Data Base and SEER, however, showed a small but significantly increased risk of death for patients younger than 45 years with lymph node metastases compared with younger patients without involved lymph nodes, and that having incrementally more metastatic lymph nodes up to six involved nodes confers additional mortality risk in this age group (340). This study underlines the importance of rigorous preoperative screening for nodal metastases and potentially raises questions about current thyroid cancer staging systems. Common to all of these studies is the conclusion that the effect of the presence or absence of lymph node metastases on overall survival, if present, is small.

The cervical node sites are well-defined (341), and the most common site of nodal metastases is in the central neck, which is cervical level VI (Fig. 3). A recent consensus conference statement describes the relevant anatomy of the central neck compartment, delineates the nodal subgroups within the central compartment commonly involved with thyroid cancer, and defines the terminology relevant to central compartment neck dissection (342). In many patients, lymph node metastases in this area do not appear abnormal on preoperative imaging (289,334,343–345) or by inspection at the time of surgery (335), defining a cN0 group.

The role of therapeutic lymph node dissection for treatment of thyroid cancer nodal metastases is well accepted for cN1 disease (336,346–348). However, the value of routine prophylactic level VI (central) neck dissection for cN0 disease remains unclear. Central compartment dissection (therapeutic or prophylactic) can be achieved with low morbidity by experienced thyroid surgeons (349–351). The value for an individual patient depends upon the utility of the staging information to the treatment team in specific patient circumstances (351,352). Based on limited and imperfect data, prophylactic dissection has been suggested to improve disease-specific survival (353), local recurrence (345,354), and post-treatment Tg levels (345,355). It has also been used to inform the use of adjuvant RAI (344,347,350,356) and improve the accuracy of the estimates of risk of recurrence (356–358). However, in several studies, prophylactic dissection has shown no improvement in long-term patient outcome, while increasing the likelihood of temporary morbidity, including hypocalcemia, although prophylactic dissection may decrease the need for repeated RAI treatments (334,346,347,349,359–364).

The removal of cN0 level VI lymph nodes detects a substantial number of patients with pN1 disease; however, the direct effect of this on long-term outcome is small at best (365,366). The use of staging information for the planning of adjuvant therapy depends upon whether this information will affect the team-based decision-making for the individual patient. For these reasons, groups may elect to include prophylactic dissection for patients with some prognostic features associated with an increased risk of metastasis and recurrence (older or very young age, larger tumor size, multifocal disease, extrathyroidal extension, known lateral node metastases) to contribute to decision-making and disease control (345,351,355). Alternatively, some groups may apply prophylactic level VI dissection to patients with better prognostic features if the patient is to have a bilateral thyroidectomy, and if the nodal staging information will be used to inform the decision regarding use of adjuvant therapy (344,350,356). Finally, for some groups it appears reasonable to use a selective approach that applies level VI lymph node dissection at the time of initial operation only to patients with clinically evident disease based on preoperative physical exam, preoperative radiographic evaluation, or intraoperative demonstration of detectable disease (cN1) (335,359,367).

The information from prophylactic central neck dissection must be used cautiously for staging information. Since microscopic nodal positivity occurs frequently, prophylactic dissection often converts patients from clinical N0 to pathologic N1a, upstaging many patients over age 45 from American Joint Committee on Cancer (AJCC) stage I to stage III (334,344–347). However, microscopic nodal positivity does not carry the recurrence risk of macroscopic clinically detectable disease (335). Thus microscopic nodal upstaging may lead to excess RAI utilization and patient follow-up. Alternatively, the demonstration of uninvolved lymph nodes by prophylactic dissection may decrease the use of RAI for some groups (344,350,356). These effects may account for some of the existing extreme variability in utilization of RAI for thyroid cancer (368).

Studies of the BRAFV600E mutation have suggested an association between presence of the mutation and the risk of nodal disease (369–371), although results across all patients with papillary thyroid carcinoma are mixed (372–375). However, the presence of a BRAFV600E mutation has a limited PPV for recurrence and therefore, BRAFV600E mutation status in the primary tumor should not impact the decision for prophylactic central neck dissection (376).
The preceding recommendations should be interpreted in light of available surgical expertise. For patients with small, noninvasive, cN0 tumors, the balance of risk and benefit may favor thyroid lobectomy and close intraoperative inspection of the central compartment, with the plan adjusted to total thyroidectomy with compartmental dissection only in the presence of involved lymph nodes.

Lymph nodes in the lateral neck (compartments II–V, Fig. 3), level VII (anterior mediastinum), and rarely in level I may also be involved by thyroid cancer (282,335,377,378). For patients in whom nodal disease is evident clinically on preoperative US and nodal FNA cytology or Tg washout measurement or at the time of surgery, surgical resection by compartmental node dissection may reduce the risk of recurrence and possibly mortality (379–381).

**RECOMMENDATION 38**

(A) Completion thyroidectomy should be offered to patients for whom a bilateral thyroidectomy would have been recommended had the diagnosis been available before the initial surgery. Therapeutic central neck lymph node dissection should be included if the lymph nodes are clinically involved. Thyroid lobectomy alone may be sufficient treatment for low-risk papillary and follicular carcinomas.

(Strong recommendation, Moderate-quality evidence)

(B) RAI ablation in lieu of completion thyroidectomy is not recommended routinely; however, it may be used to ablate the remnant lobe in selected cases.

(Weak recommendation, Low-quality evidence)

Completion thyroidectomy may be necessary when the diagnosis of malignancy is made following lobectomy for an indeterminate or nondiagnostic biopsy. In addition, some patients with malignancy may require completion thyroidectomy to provide complete resection of multicentric disease and to allow for efficient RAI therapy. However, since intrathyroidal PTC or low-risk FTC can be managed with either lobectomy or total thyroidectomy (see Recommendation 35B), a completion thyroidectomy is not always required. The surgical risks of two-stage thyroidectomy (lobectomy followed by completion thyroidectomy) are similar to those of a near-total or total thyroidectomy (382–384). The marginal utility of prophylactic lymph node dissection for cN0 disease argues against its application in re-operations.

Ablation of the remaining lobe with RAI has been used as an alternative to completion thyroidectomy (385,386). There are limited data regarding the long-term outcomes of this approach. The data suggest similar clinical outcomes with a slightly higher proportion of patients with persistent detectable Tg. This approach may be helpful in patients for whom completion thyroidectomy carries some increased risk and for whom a delay in the length of time required to achieve destruction of the normal thyroid, which follows RAI (as opposed to surgical resection), is acceptable. In one unblinded, multicenter, randomized controlled equivalence trial comparing dose activities in achieving successful ablation of a remaining lobe in patients with T1b or T2 primary tumors, who had surgical contraindications or declined completion thyroidectomy, the remnant ablation success rate was significantly higher using 100 mCi (75% success rate; 1 mCi = 37 MBq), compared with 30 mCi (54%), although mild to moderate short-term neck pain was more frequently reported in the high-dose group (66%) compared with the low-dose group (51%) (387). Prednisone treatment for neck pain was used more frequently in the high-dose group (36% of patients) than in the low-dose group.

[B10] What is the appropriate perioperative approach to voice and parathyroid issues?

[B11] Preoperative care communication

**RECOMMENDATION 39**

Prior to surgery, the surgeon should communicate with the patient regarding surgical risks, including nerve and parathyroid injury, through the informed consent process and communicate with associated physicians, including anesthesia personnel, regarding important findings elicited during the preoperative workup.

(Strong recommendation, Moderate-quality evidence)

The preoperative consent process should include explicit discussion of the potential for temporary or permanent nerve injury (and its clinical sequelae, including voice change, swallowing disability, risk of aspiration, and tracheostomy) as well as hypoparathyroidism, bleeding, scarring, disease recurrence, need for additional postoperative treatment, and need for thyroid hormone and surveillance thyroid function tests. The conversation should be informed by the operating surgeon’s own rates of complications. Results of the preoperative evaluation regarding extent of disease, risk stratification, and integrity of the airway should include results from imaging, cytology, and physical examination (388–392).

[B12] Preoperative voice assessment

**RECOMMENDATION 40**

All patients undergoing thyroid surgery should have preoperative voice assessment as part of their preoperative physical examination. This should include the patient’s description of vocal changes, as well as the physician’s assessment of voice.

(Strong recommendation, Moderate-quality evidence)

**RECOMMENDATION 41**

Preoperative laryngeal exam should be performed in all patients with

(A) Preoperative voice abnormalities

(Strong recommendation, Moderate-quality evidence)

(B) History of cervical or upper chest surgery, which places the RLN or vagus nerve at risk

(Strong recommendation, Moderate-quality evidence)

(C) Known thyroid cancer with posterior extrathyroidal extension or extensive central nodal metastases

(Strong recommendation, Low-quality evidence)
Voice alteration is an important complication of thyroid surgery affecting patients’ quality of life (with regard to voice, swallowing, and airway domains), and it can have medico-legal and cost implications (393–401).

Preoperative assessment provides a necessary baseline reference from which to establish perioperative expectations (402). Also, preoperative voice assessment may lead one to identify preoperative vocal cord paralysis or paresis, which provides presumptive evidence of invasive thyroid malignancy and is important in planning the extent of surgery and in perioperative airway management (403–405). Contralateral nerve injury at surgery in such patients could cause bilateral cord paralysis with airway implications.

Preoperative voice assessment should include the patient’s historical subjective response to questions regarding voice abnormalities or changes, as well as the physician’s objective assessment of voice, and should be documented in the medical record (Table 9) (406). Voice and laryngeal function may be further assessed through laryngoscopy, and the application of validated quality of life and auditory perceptual assessment voice instruments (402). It is important to appreciate that vocal cord paralysis, especially when chronic, may not be associated with significant vocal symptoms due to a variety of mechanisms, including contralateral vocal cord compensation. Voice assessment alone may not identify such individuals (402).

Incidence rates for preoperative vocal cord paresis or paralysis for patients with benign thyroid disease at preoperative laryngoscopy range from 0% to 3.5% and up to 8% in patients with thyroid cancer (407–411). Finding vocal cord paralysis on preoperative examination strongly suggests the presence of locally invasive disease. Approximately 10%–15% of thyroid cancers present with extrathyroidal extension, with the most common structures involved including strap muscle (53%), the RLN (47%), trachea (30%), esophagus (21%), and larynx (12%) (405,412–414).

Undiagnosed preoperative laryngeal nerve dysfunction conveys greater risk during total thyroidectomy of postoperative bilateral nerve paralysis, respiratory distress, and need for tracheostomy. Also, preoperative identification of vocal cord paralysis is important because surgical algorithms in the management of the invaded nerve incorporate nerve functional status (415).

A laryngeal exam should be performed if the voice is abnormal during preoperative evaluation. In addition, a patient should have a laryngeal exam even if the voice is normal if he or she has a history of neck surgery that placed at risk either the RLN (such as past thyroid or parathyroid surgery) or the vagus nerve (such as carotid endarterectomy, cervical esophagectomy, and anterior approach to the cervical spine) or a history of prior external beam radiation to the neck. Correlation between vocal symptoms and actual vocal cord function is poor given the potential for variation in paralytic cord position, degree of partial nerve function, and contralateral cord function/compensation; therefore, vocal symptoms may be absent in patients with vocal cord paralysis. Vocal cord paralysis may be present in 1.5% to 30% of such postsurgical patients; it can be asymptomatic in up to one-third (403,416–422).

A laryngeal exam is recommended in patients with the preoperative diagnosis of thyroid cancer if there is evidence for gross extrathyroidal extension of cancer posteriorly or extensive nodal involvement, even if the voice is normal. The laryngeal exam should be performed in the previously noted high-risk settings, but it can be performed in other patients based on the surgeon’s judgment.

**Table 9. Preoperative Factors Which May Be Associated with Laryngeal Nerve Dysfunction**

<table>
<thead>
<tr>
<th><strong>Factor</strong></th>
<th><strong>Symptoms/signs</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>History</td>
<td>Voice abnormality, dysphagia, airway symptoms, hemaoptysis, pain, rapid progression, prior operation in neck or upper chest</td>
</tr>
<tr>
<td>Physical exam</td>
<td>Extensive, firm mass fixed to the larynx or trachea</td>
</tr>
<tr>
<td>Imaging</td>
<td>Mass extending to/beyond periphery of thyroid lobe posteriorly and/or tracheoesophageal infiltration, or bulky cervical adenopathy along the course of the RLN or vagus nerve</td>
</tr>
</tbody>
</table>
with thyroid cancer) suggested statistically significant heterogeneity (variability) in treatment effect for overall and transient RLN injury, when analyzed per nerve at risk. Several studies show that intraoperative nerve monitoring is more commonly utilized by higher volume surgeons to facilitate nerve management, and several studies show improved rates of nerve paralysis with the use of neural monitoring in reoperative and complex thyroid surgery (401,426–430). Neural stimulation at the completion of lobectomy can be used as a test to determine the safety of contralateral surgery with avoidance of bilateral vocal cord paralysis, and it has been associated with a reduction of bilateral paralysis when loss of signal occurs on the first side (428,431–433). Given the complexity of monitoring systems, training and observation of existing monitoring standards are important to provide optimal benefit (424,434).

Typically, parathyroid gland preservation is optimized by gland identification via meticulous dissection (435,436). If the parathyroid(s) cannot be located, the surgeon should attempt to dissect on the thyroid capsule and ligate the inferior thyroid artery very close to the thyroid, since the majority of parathyroid glands receive their blood supply from this vessel. There are exceptions to this rule; for example, superior glands in particular may receive blood supply from the superior thyroid artery very close to the thyroid, since the majority of parathyroid glands receive their blood supply from this vessel. There are exceptions to this rule; for example, superior glands in particular may receive blood supply from the superior thyroid artery. If the parathyroid glands are inadvertently or unavoidably removed (e.g., they are intrathyroidal, or require removal during a central lymph node dissection) or devascularized, confirmation of cancer-free parathyroid tissue should be performed, and then the glands can be autotransplanted into the strap or sternocleidomastoid muscles. It is important to inspect the thyroidecomy and/or central lymphadenectomy specimen when removed and before sending it to pathology to look for parathyroid glands that can be rescued.

[RECOMMENDATION 44]

Patients should have their voice assessed in the postoperative period. Formal laryngeal exam should be performed if the voice is abnormal.

(Strong recommendation, Moderate-quality evidence)

[RECOMMENDATION 45]

Important intraoperative findings and details of postoperative care should be communicated by the surgeon to the patient and other physicians who are important in the patient’s postoperative care.

(Strong recommendation, Low-quality evidence)

Voice assessment should occur after surgery and should be based on the patient’s subjective report and physician’s objective assessment of voice in the office (409). Typically this assessment can be performed at 2 weeks to 2 months after surgery. Early detection of vocal cord motion abnormalities after thyroidecmy is important for facilitating prompt intervention (typically through early injection vocal cord medialization), which is associated with better long-term outcome, including a lower rate of formal open thyroplasty repair (437–439). Many options exist for the management of RLN paralysis, including voice therapy, vocal cord injection techniques, and open vocal cord medialization. Rates of vocal cord paralysis after thyroid surgery can only be assessed by laryngeal exam postoperatively.

Communication of intraoperative findings and postoperative care from the surgeon to other members of the patient’s thyroid cancer care team is critical to subsequent therapy and monitoring approaches. Important elements of communication include (i) surgical anatomic findings, including RLN and parathyroid status (including nerve monitoring loss of signal information if monitoring is employed); (ii) surgical disease findings, including evidence for extrathyroidal spread, completeness of tumor resection, presence and distribution of nodal disease; and (iii) postoperative status, including voice/laryngeal exam, laboratory data regarding calcium/parathyroid hormone levels and need for calcium and/or vitamin D supplementation, and the surgical pathology report (440). The surgeon should remain engaged in the patient’s pursuit care to facilitate appropriate communication and may remain engaged subsequent to endocrinologic consultation depending on regional practice patterns.

[B15] What are the basic principles of histopathologic evaluation of thyroidectomy samples?

(A) In addition to the basic tumor features required for AJCC/UICC thyroid cancer staging including status of resection margins, pathology reports should contain additional information helpful for risk assessment, such as the presence of vascular invasion and the number of invaded vessels, number of lymph nodes examined and involved with tumor, size of the largest metastatic focus to the lymph node, and presence or absence of extranodal extension of the metastatic tumor.

(Strong recommendation, Moderate-quality evidence)

(B) Histopathologic variants of thyroid carcinoma associated with more unfavorable outcomes (e.g., tall cell, columnar cell, and hobnail variants of PTC; widely invasive FTC; poorly differentiated carcinoma) or more favorable outcomes (e.g., encapsulated follicular variant of PTC without invasion, minimally invasive FTC) should be identified during histopathologic examination and reported.

(Strong recommendation, Low-quality evidence)

(C) Histopathologic variants associated with familial syndromes (cribriform-morular variant of papillary carcinoma often associated with FAP, follicular or papillary carcinoma associated with PTEN-hamartoma tumor syndrome) should be identified during histopathologic examination and reported.

(Weak recommendation, Low-quality evidence)

Pathologic examination of thyroid samples establishes the diagnosis and provides important information for risk stratification of cancer and postsurgical patient management. Histopathologically, papillary carcinoma is a well-differentiated malignant tumor of thyroid follicular cells that demonstrates characteristic microscopic nuclear features. Although a papillary growth pattern is frequently seen, it is not required for the diagnosis. Follicular carcinoma is a
well-differentiated malignant tumor of thyroid follicular cells that shows transcapsular and/or vascular invasion and lacks the diagnostic nuclear features of papillary carcinoma. Oncocytic (Hürthle cell) follicular carcinoma shows the follicular growth pattern but is composed of cells with abundant granular eosinophilic cytoplasm, which has this appearance because of accumulation of innumerable mitochondria. This tumor is currently designated by the World Health Organization as a histopathologic variant of follicular carcinoma (441). However, oncocytic follicular carcinoma tumors have some differences in biological behavior as compared to the conventional type follicular carcinoma, such as the ability to metastasize to lymph nodes and a possibly higher rate of recurrence and tumor-related mortality (269,442,443). Moreover, a growing body of genetic evidence suggests that oncocytic tumors develop via unique molecular mechanisms and therefore represent a distinct type of well-differentiated thyroid cancer (444).

Traditionally, follicular carcinomas have been subdivided into minimally invasive (encapsulated) and widely invasive. In this classification scheme, minimally invasive carcinomas are fully encapsulated tumors with microscopically identifiable foci of capsular or vascular invasion, whereas widely invasive carcinomas are tumors with extensive vascular and/or extrathyroidal invasion. More recent approaches consider encapsulated tumors with only microscopic capsular invasion as minimally invasive, whereas angioinvasive tumors are placed into a separate category (445–447). Such an approach is preferable because it distinguishes encapsulated tumors with capsular invasion and no vascular invasion, which are highly indolent tumors with a mortality ≤5%, from angioinvasive follicular carcinomas, which have a mortality ranging from 5% to 30%, depending on the number of invaded blood vessels (448).

In addition to establishing a diagnosis for each nodule in a thyroidectomy or lobectomy specimen, the pathology report must provide characteristics required for AJCC/UICC TNM staging, such as tumor size and presence of extrathyroidal extension and lymph node metastasis. Extrathyroidal extension is defined as tumor extension into the adjacent tissues. It is subdivided into minimal, which is invasion into immediate perithyroidal soft tissues or sternothyroid muscle typically detected only microscopically (T3 tumors), and extensive, which is tumor invasion into subcutaneous soft tissues, larynx, trachea, esophagus, or RLN (T4a tumors). The status of the resection (inked) margins should be reported as “involved” or “uninvolved” with tumor, since positive margins are generally associated with intermediate or high risk for recurrence.

The size of the metastatic focus in a lymph node (335) and tumor extension beyond the capsule of a lymph node (338,449,450) affect cancer risk. Therefore, the pathology report should indicate the size of the largest metastatic focus to the lymph node and the presence or absence of extranodal tumor extension, as well as the number of examined and involved lymph nodes.

Additionally, the presence of vascular (blood vessel) invasion is an unfavorable prognostic factor (451–453) and should be evaluated and reported. Vascular invasion is diagnosed as direct tumor extension into the blood vessel lumen or a tumor aggregate present within the vessel lumen, typically attached to the wall and covered by a layer of endothelial cells. More rigid criteria for vascular invasion proposed by some authors also require the presence of a fibrin thrombus attached to the intravascular tumor cells (453). The invaded blood vessels should not be located within the tumor nodule parenchyma, but rather in the tumor capsule or outside of it. Invasion of multiple (four or more) blood vessels appears to entail poorer outcomes, particularly in follicular carcinomas (454–456). Therefore, the number of invaded blood vessels (less than four or more) should be stated in the pathology report.

More than 10 microscopic variants of papillary carcinoma have been documented (457). Some of them are associated with more aggressive or conversely more indolent tumor behavior and can contribute to risk stratification. The variants with more unfavorable outcomes are the tall cell, columnar cell, and hobnail variants. The tall cell variant is characterized by predominance (≥50%) of tall columnar tumor cells whose height is at least three times their width. These tumors present at an older age and more advanced stage than classic papillary carcinoma (458–461) and demonstrate a higher recurrence rate and decreased disease-specific survival (458–460,462,463). Some studies found a higher rate of lymph node metastasis and poorer survival in patients with tall cell variant as compared to classic papillary carcinoma even in tumors without extrathyroidal extension, and this was independent of patient age and tumor size and stage (464,465). The $BRAF^{V600E}$ mutation is found in ~80% of these tumors (156,466).

The columnar cell variant of papillary carcinoma is characterized by predominance of columnar cells with pronounced nuclear stratification (467,468). These tumors have a higher risk of distant metastases and tumor-related mortality, the latter seen mostly in patients with an advanced disease stage at presentation (467–470). The $BRAF^{V600E}$ mutation is found in one-third of these tumors (467).

Papillary carcinoma with prominent hobnail features is a rare, recently described variant characterized by the predominance of cells with a hobnail appearance with apically placed nuclei and bulging of the apical cell surface (471,472). The $BRAF^{V600E}$ mutation is frequently found in these tumors (471,473). This variant of papillary carcinoma appears to be associated with frequent distant metastases (typically to lung) and increased risk of tumor-related death (471).

Other variants of papillary carcinoma, such as the solid variant and diffuse sclerosing variant, may be associated with a less favorable outcome, although the data remain conflicting. The solid variant tumors appear to be more frequently associated with distant metastases that are present in about 15% of cases, and with a slightly higher mortality rate, which was 10%–12% in two studies with 10 and 19 years mean follow-up (474,475). However, among children and adolescents with post-Chernobyl papillary carcinomas, which frequently were of the solid variant, the mortality was very low (<1%) during the first 10 years of follow-up (476,477). Importantly, the solid variant of papillary carcinoma should be distinguished from poorly differentiated thyroid carcinoma, with which it shares the insular, solid, and trabecular growth patterns. The distinction is based primarily on the preservation of nuclear features and lack of necrosis and high mitotic activity in the solid variant, as outlined by the Turin diagnostic criteria for poorly differentiated thyroid carcinoma (478). It is important to make the distinction because poorly differentiated thyroid carcinoma has a much poorer prognosis, with the 5-year survival of 72% and 10-year survival of
46% in a series of 152 patients diagnosed using the Turin criteria (479).

The prognostic implication of the diffuse sclerosing variant of papillary cancer remains controversial. This variant is characterized by diffuse involvement of the thyroid gland and a higher rate of local and distant metastases at presentation, and it has lower disease-free survival than classic papillary carcinoma (480–482). The frequency of distant metastases, predominantly affecting the lung, varies between reported series and is 10%–15% based on almost 100 published cases summarized by Lam and Lo in 2006 (483) and more recent reports. Nevertheless, the overall mortality appears to be low, with a disease-specific survival of approximately 93% at 10 years of follow-up. The diffuse sclerosing variant tends to be found in younger patients in whom response to treatment is high.

The encapsulated follicular variant of papillary carcinoma is, in contrast, associated with a low risk of recurrence, particularly in the absence of capsular or vascular invasion. This variant is characterized by a follicular growth pattern with no papillae formation and total tumor encapsulation, and the diagnosis rests on the finding of characteristic nuclear features of papillary carcinoma. Although the encapsulated follicular variant of PTC shares the follicular growth pattern with the infiltrative, nonencapsulated follicular variant of PTC, these tumors differ in their molecular profiles and biological properties. The encapsulated follicular variant tumors frequently have RAS mutations, whereas nonencapsulated follicular variants frequently harbor BRAFV600E mutations, similar to classic papillary carcinomas (484,485). Most of the encapsulated follicular variant papillary carcinomas show no invasive growth, whereas in about one-third of cases tumor capsule invasion, vascular invasion, or both are found (486,487). Whereas in the past the encapsulated follicular variant was relatively rare, at the present time half to two-thirds of all follicular variant papillary carcinomas belong to this subtype (488). The behavior of these tumors is usually quite indolent. A summary of six studies that reported 107 cases of encapsulated follicular variant revealed 25% with lymph node metastases and 1% with distant metastases (489). Among these 107 patients, one died of disease and two were alive with disease, whereas the rest (97%) of the patients were alive and well with various follow-up periods. In a study of 61 cases of encapsulated follicular variant, lymph node metastases were observed in 5%, and there was no distant metastasis (486). With median follow-up of 11 years, one patient developed tumor recurrence, and this tumor had invasion. No adverse events were found in any of the encapsulated and noninvasive tumors, including 31 patients treated with lobectomy only. Similarly, no evidence of recurrence was found in 61 out of 62 encapsulated or well-circumscribed follicular variant of PTC in another series of patients with a median follow-up of 9.2 years, and the only case that developed recurrence had a positive resection margin after initial surgery (490). In another study of a cohort of thyroid tumors followed on average for 12 years, none of 66 patients with encapsulated follicular variant of papillary carcinoma died of disease (487). Despite a low probability, some patients with encapsulated follicular variants may present with distant metastases, particularly to the bones, or develop metastasis on follow-up (491,492). Tumors prone to metastatic behavior often have a thick capsule and significant intratumoral fibrosis, and virtually all of them reveal vascular invasion or invasion of the tumor capsule. Therefore, pathologic evaluation of these tumors should include microscopic examination of the entire tumor capsule to rule out invasion, as well as careful evaluation of the tumor to rule out the presence of poorly differentiated carcinoma areas or other unfavorable diagnostic features such as tumor necrosis or high (≥3 per 10 high-power fields) mitotic activity (493). In the absence of these features, a completely excised noninvasive encapsulated follicular variant of papillary carcinoma is expected to have a very low risk of recurrence or extra-thyroidal spread, even in patients treated by lobectomy.

Similarly, excellent clinical outcomes are seen in FTCs that manifest only capsular invasion without vascular invasion (494–496). When vascular invasion is present, the tumor should no longer be designated as minimally invasive. However, some studies (456,494,497–500), although not all (496,501), suggest that only those follicular carcinomas that have a greater extent of vascular invasion (more than four foci of vascular invasion, or extracapsular vascular invasion) are associated with poorer outcomes.

Some histopathologic variants of thyroid carcinomas are important to recognize because of their association with familial tumor syndromes (41,502). The cribriform-morular variant of papillary carcinoma is frequently seen in patients with FAP due to a germline mutation in the adenomatous polyposis coli (APC) gene (503,504). It is characterized by a prominent cribriform architecture and formation of whorls or morules composed of spindle cells. The presence of aberrant β-catenin immunoreactivity provides a strong evidence for this tumor variant (505–507). Approximately 40% of patients with this variant of papillary carcinoma are found to have FAP, whereas the rest have no evidence of the inherited disease (505,508). Although no microscopic tumor features can distinguish between familial and sporadic disease, tumor multifocality is more common in the setting of the familial disease (505,508). Since many patients with the cribriform-morular variant have FAP, and thyroid cancer can precede clinically detectable colonic abnormalities in ~40% of patients (508), this diagnosis should raise the possibility of the familial disease and prompt consideration for colonic examination and genetic counseling.

Follicular carcinoma may develop as a manifestation of the PTEN hamartoma tumor syndrome, which is caused by a germline mutation in the PTEN gene (509–511). The histopathologic appearance of thyroid glands in these patients is very characteristic and should allow pathologists to suspect this syndrome (510,511). The glands typically have numerous sharply delineated, frequently encapsulated thyroid nodules that microscopically are well-delineated and cellular and have variable growth patterns (510–513). Individuals affected by this syndrome also have a high risk for benign and malignant tumors of the breast and endometrium, colon hamartomas, and others, and in light of the characteristic appearance of the thyroid gland in these patients, genetic counseling should be recommended.

Well-differentiated papillary and follicular cancers should be histologically distinguished from poorly differentiated carcinoma. Poorly differentiated carcinoma is an aggressive thyroid tumor characterized by a partial loss of the features of thyroid differentiation that occupies morphologically and behaviorally an intermediate position between well-differentiated papillary and follicular carcinomas and fully
dedifferentiated anaplastic carcinoma. Another term used in the past for this tumor was “insular carcinoma.” Diagnostic criteria for poorly differentiated carcinoma are based on the consensus Turin proposal and include the following three features: (i) solid/trabecular/insular microscopic growth pattern, (ii) lack of well-developed nuclear features of papillary carcinoma, and (iii) convoluted nuclei (evidence for partial loss of differentiation in papillary cancer), tumor necrosis, or three or more mitoses per 10 high-power fields (514). Poorly differentiated carcinomas have significantly worse outcome as compared to well-differentiated PTC and FTC, with a 10-year survival of ~50% (514–516). Patient age over 45 years, larger tumor size, presence of necrosis, and high mitotic activity are additional factors that may influence a more unfavorable outcome in patients with poorly differentiated thyroid cancer (514,517). It is not clear if the proportion of poorly differentiated carcinoma areas within the cancer nodule directly correlates with prognosis. Several studies have reported similarly decreased survival in patients with poorly differentiated carcinoma constituting more than 50% of the tumor and in those in whom it was observed as a minor component (518,519). Tumors with insular, solid, or trabecular architecture, but lacking other diagnostic features of poorly differentiated carcinoma, do not demonstrate such an aggressive behavior and therefore should not be considered as poorly differentiated. On the other hand, some studies suggest that the presence of a high mitotic rate (≥5 mitoses/10 high-power fields or Ki-67 labeling index ≥4%) or tumor necrosis predicts a less favorable outcome irrespective of the presence of the solid/trabecular/insular growth pattern (520,521).

[B16] What is the role of postoperative staging systems and risk stratification in the management of DTC?

[B17] Postoperative staging

■ RECOMMENDATION 47

AJCC/UICC staging is recommended for all patients with DTC, based on its utility in predicting disease mortality, and its requirement for cancer registries.

(Strong recommendation, Moderate-quality evidence)

Postoperative staging for thyroid cancer, as for other cancer types, is used (i) to provide prognostic information, which is of value when considering disease surveillance and therapeutic strategies, and (ii) to enable risk-stratified description of patients for communication among health care professionals, tracking by cancer registries, and research purposes.

Accurate initial staging requires a detailed understanding of all pertinent risk stratification data, whether they were obtained as part of preoperative testing, during the operation(s), or as part of postoperative follow-up. It is also important to emphasize that in many cases the written pathology report of the surgical specimen does not convey critical risk factors such as preoperative vocal cord paralysis, extent of gross extra-thyroidal invasion, completeness of resection, or remaining gross residual disease. Without these critical pieces of information, it is likely that initial risk stratification will be inaccurate and potentially misleading. Details and suggestions for effectively communicating specific risk factors between health care providers are outlined in a recent publication of the Surgical Affairs Committee of the ATA (440).

It is important to emphasize that the identification of a clinico-pathologic or molecular predictor of recurrence or mortality does not necessarily imply that more aggressive therapies (such as more extensive surgery, RAI therapy, aggressive thyroid hormone therapy with TSH suppression, targeted therapies) will have a significant impact on clinical outcomes. Similarly, the absence of a risk factor does not mean that more aggressive therapies are not indicated. Intervention studies are required to determine which at-risk-patients may benefit from additional therapies or a more conservative management approach. Until appropriate treatment intervention studies are completed, the risk stratification information associated with a clinico-pathologic risk factor or with a molecular profiling can be used as a prognostic factor to guide follow-up management decisions such as the type and frequency of imaging and biochemical testing.

[B18] AJCC/UICC TNM staging

Over the years, multiple staging systems have been developed to predict the risk of mortality in patients with DTC (522). Each of the systems uses some combination of age at diagnosis, size of the primary tumor, specific tumor histology, and extrathyroidal spread of the tumor (direct extension of the tumor outside the thyroid gland, loco-regional metastases, and/or distant metastases) to stratify patients into one of several categories with differing risks of death from thyroid cancer. Recently, a nomogram was developed and validated using the SEER data base, which provides a mathematical approach to integrating these important clinical predictive features into a specific mortality risk estimate for an individual patient (523). Using an approach similar to the MACIS system from the Mayo Clinic (270), a quantitative approach based on histology, age, lymph node metastases, tumor size, and extrathyroidal extension utilizing TNM staging has recently been proposed and validated (524,525).

While none of the staging systems has been shown to be clearly superior to the other systems, several studies have demonstrated that the AJCC/UICC TNM system (Table 10) and the MACIS system consistently provide the highest proportion of variance explained (PVE, a statistical measure of how well a staging system can predict the outcome of interest) when applied to a broad range of patient cohorts (277,501,526–530), and they have been validated in retrospective studies as well as prospectively in clinical practice. Unfortunately, none of the staging systems designed to predict mortality from thyroid cancer can account for more than a small proportion (5%–30%, corresponding to PVE values of 0.05–0.30) of the uncertainty associated with eventual death from thyroid cancer (277,501,526–530). This relative inability to accurately predict the risk of death from thyroid cancer for an individual patient may be related to the failure of current staging systems to adequately integrate the risk associated with other potentially important clinico-pathologic features such as the specific histology (well-differentiated thyroid cancer versus poorly differentiated thyroid cancer), molecular profile, size and location of distant metastases (pulmonary metastases versus bone metastases versus brain metastases), functional status of the metastases (RAI avid versus 18F-FDG-PET avid), and effectiveness of initial therapy (completeness of resection, effectiveness of
### Table 10. AJCC 7th Edition/TNM Classification System for Differentiated Thyroid Carcinoma

<table>
<thead>
<tr>
<th>Stage</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>T0</td>
<td>No evidence of primary tumor</td>
</tr>
<tr>
<td>T1a</td>
<td>Tumor ≤1 cm, without extrathyroidal extension</td>
</tr>
<tr>
<td>T1b</td>
<td>Tumor &gt;1 cm but ≤2 cm in greatest dimension, without extrathyroidal extension</td>
</tr>
<tr>
<td>T2</td>
<td>Tumor &gt;2 cm but ≤4 cm in greatest dimension, without extrathyroidal extension</td>
</tr>
<tr>
<td>T3</td>
<td>Tumor &gt;4 cm in greatest dimension limited to the thyroid or Any size tumor with minimal extrathyroidal extension (e.g., extension into sternothyroid muscle or perithyroidal soft tissues)</td>
</tr>
<tr>
<td>T4a</td>
<td>Tumor of any size extending beyond the thyroid capsule to invade subcutaneous soft tissues, larynx, trachea, esophagus, or recurrent laryngeal nerve</td>
</tr>
<tr>
<td>T4b</td>
<td>Tumor of any size invading prevertebral fascia or encasing carotid artery or mediastinal vessels</td>
</tr>
<tr>
<td>N0</td>
<td>No metastatic nodes</td>
</tr>
<tr>
<td>N1a</td>
<td>Metastases to level VI (pretracheal, paratracheal, and prelaryngeal/Delphian lymph nodes)</td>
</tr>
<tr>
<td>N1b</td>
<td>Metastases to unilateral, bilateral, or contralateral cervical (levels I, II, III, IV, or V) or retropharyngeal or superior mediastinal lymph nodes (level VII)</td>
</tr>
<tr>
<td>M0</td>
<td>No distant metastases</td>
</tr>
<tr>
<td>M1</td>
<td>Distant metastases</td>
</tr>
</tbody>
</table>

**Patient age <45 years old at diagnosis**

<table>
<thead>
<tr>
<th>Stage</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>Any T Any N M0</td>
</tr>
<tr>
<td>II</td>
<td>Any T Any N M1</td>
</tr>
</tbody>
</table>

**Patient age ≥45 years old at diagnosis**

<table>
<thead>
<tr>
<th>Stage</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>T1a N0 M0</td>
</tr>
<tr>
<td>I</td>
<td>T1b N0 M0</td>
</tr>
<tr>
<td>II</td>
<td>T2 N0 M0</td>
</tr>
<tr>
<td>III</td>
<td>T1a N1a M0</td>
</tr>
<tr>
<td>III</td>
<td>T1b N1a M0</td>
</tr>
<tr>
<td>III</td>
<td>T2 N1a M0</td>
</tr>
<tr>
<td>III</td>
<td>T3 N0 M0</td>
</tr>
<tr>
<td>III</td>
<td>T3 N1a M0</td>
</tr>
<tr>
<td>IVa</td>
<td>T1a N1b M0</td>
</tr>
<tr>
<td>IVa</td>
<td>T1b N1b M0</td>
</tr>
<tr>
<td>IVa</td>
<td>T2 N1b M0</td>
</tr>
<tr>
<td>IVa</td>
<td>T3 N1b M0</td>
</tr>
<tr>
<td>IVa</td>
<td>T4a N0 M0</td>
</tr>
<tr>
<td>IVa</td>
<td>T4a N1a M0</td>
</tr>
<tr>
<td>IVa</td>
<td>T4a N1b M0</td>
</tr>
<tr>
<td>IVb</td>
<td>T4b Any N M0</td>
</tr>
<tr>
<td>IVc</td>
<td>Any T Any N M1</td>
</tr>
</tbody>
</table>

Used with the permission of the American Joint Committee on Cancer (AJCC), Chicago, Illinois. The original source for this material is the AJCC Cancer Staging Manual, Seventh Edition (1077) published by Springer Science and Business Media LLC (http://www.springer.com).

**Table Notes**

- **RAI** external beam radiation therapy or other systemic therapies. Furthermore, recent studies have questioned the use of the age of 45 years as a cutoff to upstage patients using the AJCC/UICC TNM system (340,531–533).
- Even though the various staging systems designed to predict mortality from thyroid cancer were developed and validated using cohorts that were either exclusively or predominantly PTC patients, several small studies have demonstrated that MACIS and AJCC/UICC TNM staging systems were also predictive in patients with FTC (501,534,535).
- Currently, none of the mortality risk systems incorporate molecular testing results. This may need to be re-evaluated as studies emerge using molecular testing including BRAFV600E, TERT, and TP53 or combinations of markers. For example, in one study that analyzed more than 400 DTCs, the presence of a TERT mutation was found to be an independent predictor of mortality (hazard ratio [HR] 10.35 [95% CI 2.01–53.24]) for all differentiated cancers and for papillary carcinomas (154). These potential prognostic markers are promising, but require further study.

**[B19] What initial stratification system should be used to estimate the risk of persistent/recurrent disease?**

**RECOMMENDATION 48**

(A) The 2009 ATA Initial Risk Stratification System is recommended for DTC patients treated with thyroidectomy, based on its utility in predicting risk of disease recurrence and/or persistence.

**(Strong recommendation, Moderate-quality evidence)**

(B) Additional prognostic variables (such as the extent of lymph node involvement, mutational status, and/or the degree of vascular invasion in FTC), not included in the 2009 ATA Initial Risk Stratification system may be used to further refine risk stratification for DTC as described in the following text (and in Fig. 4) in the Modified Initial Risk Stratification system. However, the incremental benefit of adding these specific prognostic variables to the 2009 Initial Risk Stratification system has not been established.

**(Weak recommendation, Low-quality evidence)**

(C) While not routinely recommended for initial postoperative risk stratification in DTC, the mutational status of BRAF, and potentially other mutations such as TERT, have the potential to refine risk estimates when interpreted in the context of other clinico-pathologic risk factors.

**(Weak recommendation, Moderate-quality evidence)**

Because the AJCC/TNM risk of mortality staging system does not adequately predict the risk of recurrence in DTC (536–539), the 2009 version of the ATA thyroid cancer guidelines proposed a three-tiered clinico-pathologic risk stratification system that classified patients as having low, intermediate, or high risk of recurrence (25). Low-risk patients were defined as having intrathyroidal DTC with no evidence of extrathyroidal extension, vascular invasion, or metastases. Intermediate-risk patients demonstrated either microscopic extrathyroidal extension, cervical lymph node metastases, RAI-avid disease in the neck outside the thyroid capsule, or evidence of extrathyroidal extension, vascular invasion, or metastases. High-risk patients demonstrated evidence of extrathyroidal extension, vascular invasion, or metastases.
bed, vascular invasion, or aggressive tumor histology. High-risk patients had gross extrathyroidal extension, incomplete tumor resection, distant metastases, or inappropriate postoperative serum Tg values (Table 11).

The 2009 ATA risk stratification system was somewhat different than staging systems proposed by a European Consensus conference (540) and the Latin American Thyroid Society (LATS) (541) which classify patients as either being at very low risk (unifocal, intrathyroidal T1aN0M0), low risk (T1b N0M0, T2N0M0, or multifocal T1N0M0), or high risk (any T3 or T4, any N1, or any M1). The European and LATS very low risk and low risk categories would be classified as ATA low risk, while the ETA high risk category would be subdivided between ATA intermediate risk (minor extrathyroidal extension, N1 disease) and ATA high risk (gross extrathyroidal extension, M1, incomplete tumor resection).

Subsequent studies have retrospectively validated the 2009 ATA risk of recurrence staging system through analysis of independent datasets originating from three respective continents (Table 12). These studies have reported the estimates of patients who subsequently had no evidence of disease (NED) in each ATA Risk Category after total thyroidectomy and RAI remnant ablation: (a) low risk, 78%–91% NED, (b) intermediate risk, 52%–64% NED, and (c) high risk, 31%–32% NED (538,539,542,543). In these datasets, NED was defined as a stimulated Tg <1 ng/mL with no other radiological or clinical evidence of disease. Prospectively collected validation data for the ATA initial risk stratification system are needed.

Three additional studies, in which the ATA risk classification system was retrospectively evaluated, have also suggested that the ATA risk of recurrence model may be applied in low- and intermediate-risk patients in the absence of RAI ablation (328,544,545). Over a median follow-up period that ranged from 5 to 10 years, structural disease recurrence was identified in less than 1%–2% of ATA low-risk patients and 8% of ATA intermediate-risk patients who underwent thyroid surgery without RAI ablation as the initial therapy (328,544,545).

The type of persistent disease also varies according to ATA initial risk stratification, with 70%–80% of the persistent disease in ATA low-risk patients manifested by abnormal serum Tg levels (suppressed or stimulated Tg >1 ng/mL) without structurally identifiable disease, while only 29%–51% of the ATA intermediate-risk patients and 19%–21% of ATA high-risk patients are classified as having persistent disease only on the basis of biochemical abnormalities (538,539,543). With increasing ATA risk level, the RR of having structural persistent/recurrent disease increases. Thus, the ATA low-risk patients appear to have the highest RR of
isolated thyroglobulinemia, which may be of less clinical significance than structural disease persistence or recurrence. Similar to what was seen with the staging systems designed to predict risk of mortality from thyroid cancer (see section [B18] above), the PVE by the ATA risk of recurrence system was suboptimal, ranging from 19% to 34% (538,542). More recently, a novel mathematical clinico-pathologic staging system from the University of Yonsei yielded similar results with a PVE of 11.9% in predicting disease recurrence (546).

**Potential impact of specific clinico-pathologic features on the risk estimates in PTC**

As originally conceived, the ATA initial risk stratification system was a three-tiered system in which clinico-pathologic features available at the time of initial treatment were used to classify DTC patients as having either low, intermediate, or high risk of either recurrence or persistent disease. However, as with any categorical staging system, the risk of recurrence within the individual risk categories (low, intermediate, and high) is not as accurate as might be desired.

### Table 11. ATA 2009 Risk Stratification System with Proposed Modifications

<table>
<thead>
<tr>
<th>ATA low risk</th>
<th>Papillary thyroid cancer (with all of the following):</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>• No local or distant metastases;</td>
</tr>
<tr>
<td></td>
<td>• All macroscopic tumor has been resected</td>
</tr>
<tr>
<td></td>
<td>• No tumor invasion of loco-regional tissues or structures</td>
</tr>
<tr>
<td></td>
<td>• The tumor does not have aggressive histology (e.g., tall cell, hobnail variant, columnar cell carcinoma)</td>
</tr>
<tr>
<td></td>
<td>• If 131I is given, there are no RAI-avid metastatic foci outside the thyroid bed on the first posttreatment whole-body RAI scan</td>
</tr>
<tr>
<td></td>
<td>• No vascular invasion</td>
</tr>
<tr>
<td></td>
<td>• Clinical N0 or ≤5 pathologic N1 micrometastes (&lt;0.2 cm in largest dimension)¹</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ATA intermediate risk</th>
<th>Microscopic invasion of tumor into the perithyroidal soft tissues</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RAI-avid metastatic foci in the neck on the first posttreatment whole-body RAI scan</td>
</tr>
<tr>
<td></td>
<td>Aggressive histology (e.g., tall cell, hobnail variant, columnar cell carcinoma)</td>
</tr>
<tr>
<td></td>
<td>Papillary thyroid cancer with vascular invasion</td>
</tr>
<tr>
<td></td>
<td>Clinical N1 or &gt;5 pathologic N1 with all involved lymph nodes &lt;3 cm in largest dimension¹</td>
</tr>
<tr>
<td></td>
<td>Multifocal papillary microcarcinoma with ETE and ( BRAF^{V600E} ) mutated (if known)⁶</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ATA high risk</th>
<th>Macroscopic invasion of tumor into the perithyroidal soft tissues (gross ETE)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Incomplete tumor resection</td>
</tr>
<tr>
<td></td>
<td>Distant metastases</td>
</tr>
<tr>
<td></td>
<td>Postoperative serum thyroglobulin suggestive of distant metastases</td>
</tr>
<tr>
<td></td>
<td>Pathologic N1 with any metastatic lymph node ≥3 cm in largest dimension¹</td>
</tr>
<tr>
<td></td>
<td>Follicular thyroid cancer with extensive vascular invasion (&gt;4 foci of vascular invasion)¹</td>
</tr>
</tbody>
</table>

¹Proposed modifications, not present in the original 2009 initial risk stratification system. See sections [B19]–[B23] and Recommendation 48B.

### Table 12. American Thyroid Association Risk Stratification System: Clinical Outcomes Following Total Thyroidectomy and Radioiodine Remnant Ablation or Adjuvant Therapy

<table>
<thead>
<tr>
<th>ATA risk</th>
<th>Study</th>
<th>NED, %</th>
<th>Biochemical incomplete, %b</th>
<th>Structural incomplete, %c</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>Tuttle et al. (538)</td>
<td>86</td>
<td>11</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Castagna et al. (542)</td>
<td>91</td>
<td>ND¹</td>
<td>ND¹</td>
</tr>
<tr>
<td></td>
<td>Vaisman et al. (539)</td>
<td>88</td>
<td>10</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Pitoia et al. (543)</td>
<td>78</td>
<td>15</td>
<td>7</td>
</tr>
<tr>
<td>Intermediate</td>
<td>Tuttle et al. (538)</td>
<td>57</td>
<td>22</td>
<td>21</td>
</tr>
<tr>
<td></td>
<td>Vaisman et al. (539)</td>
<td>63</td>
<td>16</td>
<td>21</td>
</tr>
<tr>
<td></td>
<td>Pitoia et al. (543)</td>
<td>52</td>
<td>14</td>
<td>34</td>
</tr>
<tr>
<td>High</td>
<td>Tuttle et al. (538)</td>
<td>14</td>
<td>14</td>
<td>72</td>
</tr>
<tr>
<td></td>
<td>Vaisman et al. (539)</td>
<td>16</td>
<td>12</td>
<td>72</td>
</tr>
<tr>
<td></td>
<td>Pitoia et al. (543)</td>
<td>31</td>
<td>13</td>
<td>56</td>
</tr>
</tbody>
</table>

¹Because the ATA intermediate- and high-risk groups were merged into a single “high-risk” group in the series by Castagna et al. (542), risk of persistent/recurrent disease for these subgroups is not presented.

²Proportion of patients with a biochemical incomplete response. Definition: suppressed Tg >1 ng/mL, TSH-stimulated Tg >10 ng/mL, or rising anti-Tg antibody levels in the absence of structural disease.

³Proportion of patients with persistent/recurrent disease that is structural. Definition: structural disease that is either biopsy-proven or highly suspicious for disease with or without abnormal serum Tg.

ND, not determined.
high) can vary depending on the specific clinical features of individual patients (Fig. 4). In addition, the three-tiered system did not specifically address the risk of recurrence associated with specific DTC histologies, multifocality, genotype, extent of vascular invasion, or extent of metastatic lymph node involvement.

For example, while intrathyroidal PTCs of all sizes are included in the ATA low-risk category, the risk of structural disease recurrence can vary from 1%–2% in unifocal papillary microcarcinomas, to 4%–6% in multifocal papillary microcarcinomas (141,147), to 5%–6% in 2–4 cm intrathyroidal PTC (547), and 8%–10% in intrathyroidal PTC >4 cm (547). Similarly, while all DTC patients with locoregional lymph node metastases were classified as having intermediate risk in the 2009 ATA risk stratification system, the risk of structural disease recurrence can vary from 4% in patients with fewer than five metastatic lymph nodes, to 5% if all involved lymph nodes are <0.2 cm, to 19% if more than five lymph nodes are involved, to 21% if more than 10 lymph nodes are involved, to 22% if macroscopic lymph node metastases are clinically evident (clinical N1 disease), and 27%–32% if any metastatic lymph node is >3 cm (335,548). These risk estimates apply to both N1a and N1b disease because there are insufficient data to determine risk based on location within the neck independent of size, number of involved nodes, and extranodal extension. It is very likely that these risk estimates could vary depending on the types and extent of nodal dissections done in individual patients. As defined by the ATA surgical affairs committee task force on thyroid cancer nodal surgery, ≤5 pN1 micrometastases (<0.2 cm in largest dimension) are classified as lower risk N1 disease (<5% risk of recurrence) (335). Clinical N1 disease, more than five metastatic lymph node, or any metastatic lymph node >3 cm in largest dimension are classified as higher risk N1 disease (>20% risk of recurrence) (335). Identification of extranodal extension of the tumor through the metastatic lymph node capsule has also been associated with an increased risk of recurrent/persistent disease (338,450). It is difficult to estimate the risk associated with extension of the tumor through the capsule of involved lymph nodes because this histologic finding is tightly linked with both the number of involved lymph nodes (338,450) and invasion of the tumor through the thyroid capsule (450,549).

It is also important to differentiate the clinical significance of minor extrathyroidal extension (ATA intermediate risk) from gross extrathyroidal invasion of surrounding structures (ATA high risk). The risk of recurrence associated with minor extrathyroidal extension (pT3 disease manifest by minimal extrathyroidal extension) ranges from 3% to 9% (350–554), while the risk of recurrence in patients with gross extrathyroidal extension (pT4a disease involving the subcutaneous soft tissues, larynx, trachea, esophagus, or RLN) ranges from 23% to 40% (537,550,552–555).

The encapsulated follicular variant of papillary thyroid carcinoma also appears to be associated with a low 10-year risk of recurrence (486,490,556). Only two recurrences were reported in the 152 patients (1.3% risk of recurrence) described in these three reports. No recurrences were described in the 42 patients that had encapsulated FVPTC without capsular or lymphovascular invasion in the series by Liu et al. (486). However, the study by Baloch et al. (556) included a single patient with encapsulated FVPTC that developed bone metastasis despite not having lymphovascular invasion or lymph node metastases identified at the time of diagnosis. Therefore, intrathyroidal FVPTC is best classified as a low-risk tumor that is unlikely to recur or metastasize.

Well-differentiated PTCs demonstrating only capsular invasion (without vascular invasion) usually have an excellent prognosis with recurrence rates of 0%–7% and can be classified as low-risk tumors (494–496). Encapsulated, minimally invasive PTC with only minor vascular invasion (small number of foci confined to intracapsular vessels) also appears to have a low recurrence rate of approximately 0%–5% (456,500). Furthermore, some studies (456,494,497–500), but not all (496,501), suggest a greater extent of vascular invasion (more than four foci of vascular invasion, or extracapsular vascular invasion) is associated with poorer outcomes even in encapsulated PTCs. However, FTC is considered ATA high risk if extensive vascular invasion is present because the risk of the development of newly identified distant metastases is as high as 30%–55% (455,494,497,500).

In PTC, most (451–453,558–561), but not all (562) studies demonstrate that vascular invasion is associated with worse clinical outcomes. Recurrence rates were significantly higher if vascular invasion was present in the studies by Gardner et al. (16%–20% with vascular invasion, 3%–6% without), Nishida et al. (28% with vascular invasion, 15% without), and Falvo et al. (30% with vascular invasion, 5% without), but not in the studies by Furlan et al. (562) or Akslen et al. (559). Vascular invasion in PTC was also associated with higher rates of distant metastases (453,561) and disease-specific mortality (559,560).

[B21] Potential impact of BRAFV600E and other mutations on risk estimates in PTC

In a pooled univariate analysis of 1849 PTC patients, the presence of a BRAFV600E mutation was associated with increased disease-specific mortality, although this was not significantly associated with mortality in a multivariate analysis (370). However, a significant interaction between BRAFV600E mutation and several conventional clinico-pathological risk factors was seen, such that the risk of mortality was higher in patients with BRAF mutation compared to those with wild-type BRAF in the setting of lymph node metastases (11.1% vs. 2.6%, p < 0.001), distant metastases (51.5% vs. 18.2%, p < 0.001), AJCC stage IV disease (31.4% vs. 13%, p = 0.004), and age ≥45 years at diagnosis (8% vs. 1.9%, p < 0.001). In a systematic review and meta-analysis of 14 publications that included 2470 PTC patients from nine different countries, the BRAFV600E mutation was associated with a significantly higher risk of recurrence than BRAF wild-type tumors (24.9% vs. 12.6%, p < 0.00001 [95% CI 1.61–2.32]) (563). In the studies included in this meta-analysis, the risk of recurrence in BRAFV600E–positive tumors ranged from 11% to 40% (median 26.5%), while the risk of recurrence in BRAF wild-type tumors ranged from 2% to 36% (median 9.5%). Because the BRAFV600E mutation is tightly linked with the presence of aggressive histologic phenotypes, lymph node metastases, and extrathyroidal extension, it is difficult to determine the proportion of risk that is attributable to the BRAF mutation versus that attributable to the other clinico-pathologic features. Some studies (564,565), but not all (566–568), have demonstrated that the BRAFV600E mutation is an independent predictor of risk of
recurrence in multivariate analysis. Furthermore, a recent publication demonstrated a small, but statistically significant improvement in risk stratification if *BRAF* status was used in conjunction with the 2009 ATA initial risk stratification system (569). In a meta-analysis of 2167 patients, the presence of *BRAF*V600E mutation had a sensitivity of 65% in identifying those tumors that subsequently recurred, but had a PPV of only 25% in predicting the risk of recurrence (563). Unlike previous studies that were often small and primarily single-institution reports, a multicenter pooled-data evaluation of 2099 patients (1615 women, 484 men, median age 45 years) demonstrated that the *BRAF*V600E mutation was significantly associated with the risk of recurrence in both classical PTC (20.7% vs. 12.4%, HR 1.46 [95% CI 1.08–1.99] after adjustment for multiple clinico-pathological features), and FVPTC (21.3% vs. 7.0%, HR 3.20 [95% CI 1.46–7.02] after adjustment for multiple clinico-pathological features). Based on these data, it appears that the *BRAF* status in isolation is not sufficient to substantially contribute to risk stratification in most patients. However, an incremental improvement in risk stratification can be achieved if the *BRAF* mutational status is considered in the context of other standard clinico-pathological risk factors. Since the clinical implications of this incremental improvement in risk stratification are not clear, we are not routinely recommending *BRAF* mutational evaluation for initial post-operative risk stratification in DTC.

The potential role of *BRAF* status in isolation as an aid to risk stratification in patients clinico-pathologically classified as having ATA low risk is currently being evaluated. In a cohort of low-risk patients with intrathyroidal PTC (<4 cm, N0, M0; 33% with *BRAF* mutation), the overall risk of having structural disease recurrence over 5 years of follow up was 3% (565). However, *BRAF*V600E-mutated tumors had a recurrence rate of 8% (8 of 106) compared with only 1% (2 of 213) in *BRAF* wild-type tumors (*p* = 0.003, Fisher’s exact). Furthermore, in multivariate analysis, the only clinico-pathological significant predictor of persistent disease after 5 years of follow-up was the presence of a *BRAF*V600E mutation. If these findings are verified in additional studies, it is possible that *BRAF* testing could be used to help further risk stratify patients with intrathyroidal PTC as having very low risk of recurrence (*BRAF* wild-type) or intermediate risk of recurrence (*BRAF*V600E mutation).

The impact of *BRAF* status on the risk of recurrence in the very low-risk patients (intrathyroidal unifocal papillary microcarcinomas <1 cm) appears to be small. Even though *BRAF* mutation is present in 30%–67% of papillary microcarcinomas (152,153,156,566,570–576), the overall clinical recurrence rate is quite low, ranging from 1% to 6% (274,566). In a series of 99 papillary microcarcinoma patients with an overall recurrence rate of 7%, no recurrences were detected in the patients with a *BRAF*V600E mutation and intrathyroidal, unifocal tumors. Conversely, *BRAF*V600E-mutated multifocal PTMC with extrathyroidal extension demonstrated a 20% recurrence rate (150). Therefore, in the absence of data demonstrating that the *BRAF*V600E mutation is associated with increased structural recurrence in very low-risk tumors, we have classified intrathyroidal papillary microcarcinomas (T1a, N0, M0) harboring *BRAF*V600E mutations with no other worrisome features (such as extrathyroidal extension, aggressive histology, vascular invasion, or lymph node metastases) as ATA low-risk tumors. The few patients (about 10% of PTMC patients) that demonstrate multifocal PTMC with extrathyroidal invasion and a *BRAF*V600E mutation would be considered as having ATA intermediate risk for recurrence. Based on these data, there appears to be little role for *BRAF* mutational testing as an aid to risk stratification in PTMC tumors that do not demonstrate other worrisome clinico-pathologic features.

More recent data suggest that aggressive behavior of a given thyroid carcinoma, including high probability of tumor recurrence, is likely when it harbors more than one known oncogenic mutation, and specifically a *BRAF* mutation co-occurring with a *TERT* promoter, *PIK3CA*, *TP53*, or *AKT1* mutation (155–157,577). Such a combination of several mutations is seen in a much smaller fraction of PTC as compared with a 40%–45% incidence of *BRAF* mutations and is expected to serve as a more specific marker of unfavorable outcomes of PTC.

Two other molecular markers that appear to confer an increased risk of tumor recurrence and tumor-related mortality are *TP53* and *TERT* mutations. *TP53* mutations have been known to occur mostly in poorly differentiated and anaplastic thyroid cancers. However, more recent broad mutational analyses identified *TP53* mutations in 2 of 57 (3.5%) well-differentiated PTC and 4 of 36 (11%) of well-differentiated FTC (578). Both PTCs in this series that were positive for *TP53* mutations also showed mutations in *BRAF* (or *BRAF* and *PIK3CA*) and developed lung metastases. All four *TP53*-positive FTC (with no other coexisting mutations) were oncocytic, and three out of four of those were widely invasive FTC.

Finally, recent studies identified *TERT* promoter mutations as a likely predictor of more unfavorable outcomes for patients with thyroid cancer. *TERT* mutations were found in 7%–22% of PTC and 14%–17% of FTC, but with a significantly higher prevalence in dedifferentiated thyroid cancers (154,579–581). In some (154,579), but not all (580) reports, *TERT* mutations were found more often in PTC carrying a *BRAF* mutation. In the largest reported series (332 PTC and 70 FTC followed on average for 8 years), a *TERT* mutation was an independent predictor of disease-free survival (odds ratio [OR] 4.08 [95% CI 1.54–14.27]) and mortality (HR 10.35 [95% CI 2.01–53.24]) for well-differentiated thyroid cancer (154). Furthermore, the combination of a *TERT* mutation and a *BRAF* mutation within the same tumor was associated with a high risk of structural disease recurrence (155). These results, although pending confirmation in other studies, suggest that these molecular markers, alone or in combination, may be helpful for risk stratification of thyroid cancer and provide significantly more accurate risk assessment than *BRAF* mutational status taken in isolation.

[B22] Potential impact of postoperative serum Tg on risk estimates

Several studies have demonstrated the clinical utility of a serum Tg measurement (either TSH stimulated or non-stimulated) obtained a few weeks after total thyroidectomy (postoperative Tg) and before RAI remnant ablation as a tool to aid in initial risk stratification and adjuvant therapy decision-making (See Recommendations 50B and 50C). Please see section [C6] for discussion of Tg measurements.
Proposed modifications to the 2009 ATA initial risk stratification system

While the 2009 risk stratification system has proven to be a valuable tool for initial risk stratification in PTC, modifications are required to better incorporate our new understanding regarding the risks associated with the extent of lymph node involvement, mutational status, and specific PTC histologies (Table 11). While the modified 2009 risk stratification system continues to classify intrathyroidal PTC without vascular invasion as low risk, the category was expanded to include patients with small-volume lymph node metastases (clinical N0 or ≤5 pathologic N1 micrometastases, <0.2 cm in largest dimension), intrathyroidal encapsulated follicular variant of PTC, intrathyroidal well-differentiated follicular cancer with capsular or minor vascular invasion (<4 vessels involved), and intrathyroidal papillary microcarcinomas that are either BRAF wild-type or BRAF mutated. Similarly, the modified 2009 intermediate-risk category continues to include patients with microscopic invasion of the tumor into perithyroidal soft tissues, vascular invasion, uptake outside the thyroid bed at the time of remnant ablation, and aggressive histologies, but it has been modified to include only a subset of patients with lymph node metastases (clinical N1 or >5 pathologic N1 with all involved lymph nodes ≤3 cm in largest dimension and multifocal papillary microcarcinoma with extrathyroidal extension and BRAF mutated (if known)). Finally, the high-risk category continues to include patients with macroscopic extrathyroidal extension, incomplete tumor resection, distant metastases, and postoperative serum Tg suggestive of distant metastases, but it has been expanded to include patients with large-volume lymph node involvement (any metastatic lymph node ≥3 cm in largest dimension), and FTC with extensive vascular invasion (>4 foci of vascular invasion or extracapsular vascular invasion). While these modifications are based on our current literature review of the impact of the individual risk factors, future studies will be required to validate this proposed modification to determine if the additional modifying factors provide significant incremental improvement in risk stratification.

Risk of recurrence as a continuum of risk

While the ATA initial risk stratification system provides a meaningful and valuable tool for predicting risk of recurrence when used as a three-tiered categorical staging system, additional insights can be gained if one appreciates that the risk of structural disease recurrence is a continuum of risk that ranges from <1% in very low-risk patients to >50% in high-risk patients (see Fig. 4). Therefore, individualized management recommendations should be based not only on the categorical risk of recurrence estimate, but also on a more individualized estimate of risk in which the ATA low- and intermediate-risk categories are further characterized on the basis of respective clinical or pathologic features, such as the size of the primary tumor, number and size of loco-regional lymph node metastases (as well as extranodal extension), specific histologic variant, vascular invasion, extent of extrathyroidal extension, or other potentially important factors.

While the risk estimates presented in this section may be useful in guiding initial treatment selections (extent of initial surgery, need for RAI ablation), it is important to recognize that these risk estimates likely reflect not only the tumor biology, but also the impact of initial therapy, which varies from thyroid lobectomy to total thyroidectomy with varying extents of lymph node dissection, either with or without RAI ablation, across the various studies. Additional studies are needed to better define the risk of recurrence in most of these clinical scenarios in patients randomized to receive one of several initial treatment options (lobectomy versus total thyroidectomy, extent of lymph node dissection, with or without RAI ablation).

How should initial risk estimates be modified over time?

RECOMMENDATION 49

Initial recurrence risk estimates should be continually modified during follow-up, because the risk of recurrence and disease-specific mortality can change over time as a function of the clinical course of the disease and the response to therapy.

(Strong recommendation, Low-quality evidence)

While initial staging systems provide important insights into an individual patient’s risk of recurrence and disease-specific mortality, they provide static, single-point estimates of risk based only on data available at the time of initial therapy. None of the currently available initial staging systems are capable of using new data obtained during the course of follow-up to modify the initial risk estimate. For example, an ATA low-risk patient that demonstrates a rising serum Tg associated with cervical lymphadenopathy highly suspicious for recurrent disease at some point during follow-up would still be classified as ATA low risk despite the presence of clinical data demonstrating a high risk of recurrence. Conversely, an ATA high-risk patient that remains NED over 30 years of appropriate follow-up would still be classified as ATA high risk despite having a risk of recurrence that is significantly lower than would have been predicted at the time of initial therapy.

Therefore, while the initial staging systems can be informative in guiding therapeutic and early diagnostic follow-up strategy decisions, a risk stratification system that incorporates individual response to therapy into a real-time, dynamic risk stratification scheme is needed to provide an individualized approach to ongoing management (582,583). One approach that has been proposed is to use the risk estimates from the initial staging systems to guide initial management recommendations and then to incorporate a response-to-therapy assessment during follow-up to modify these initial risk estimates in an ongoing, dynamic process (584).

Multiple studies have now shown that many patients initially classified as intermediate or high risk of recurrence using initial staging systems can be reclassified as having a subsequent low risk of recurrence based on having an excellent response to initial therapy (538,539,542,586–601). Furthermore, the PVE values associated with staging systems that incorporated response-to-therapy variables into revised risk estimates were significantly higher (62%–84%) than those seen with initial staging systems (<30%) (538,542). These data indicate that long-term outcomes can be more reliably predicted using systems that adjust to new data over time.
[B26] Proposed terminology to classify response to therapy and clinical implications

All clinical, biochemical, imaging (structural and functional), and cytopathologic findings obtained during follow-up should be used to redefine the clinical status of the patient and to assess their individual response to therapy. Ideally, the global consideration of the composite findings of such an assessment would allow for the disease status to be classified into one of several discrete response to therapy risk strata as proposed by Tuttle et al. (538,582). Potential challenges in applying this specific system in routine clinical practice include lack of validation in specific subgroups of patients (such as those who had less than total thyroidec- tomy and RAI ablation), the lack of published prospective data utilizing this system in clinical care, and some inconsistency with other authors in classifying the significance of varying levels of detectable Tg levels or imaging findings.

The response to therapy restaging system was also not designed or specifically tested by developers to guide specific therapeutic decisions on primary therapy (such as use of adjuvant treatment) because it has been designed for use after primary therapy is completed. Prospective studies of the value of this system for guiding extent of primary treatment, including adjuvant treatment decisions, are needed. However, given that there is emerging evidence that such a reclassification system has potential to be of great importance in ongoing clinical care of DTC patients after primary treatment, the details are described herein.

It is acknowledged that appropriate clinical application of such a system is highly dependent on the availability of high-quality biochemical testing and structural, and functional imaging with appropriate interpretation (e.g., standardized radiologic reporting). Therefore, applicability may be limited in settings where these procedures are not available or feasible.

The concept and initial validation of the four response-to-therapy categories presented here were described by Tuttle et al. (538) and modified in Vaisman et al. (328). As originally conceived, these clinical outcomes described the best response to initial therapy during the first 2 years of follow-up (538,582), but they are now being used to describe the clinical status at any point during follow-up.

- **Excellent response**: no clinical, biochemical, or structural evidence of disease
- **Biochemical incomplete response**: abnormal Tg or rising anti-Tg antibody levels in the absence of localizable disease. Please see section [C6] for discussion of Tg measurements.
- **Structural incomplete response**: persistent or newly identified loco-regional or distant metastases
- **Indeterminate response**: nonspecific biochemical or structural findings that cannot be confidently classified as either benign or malignant. This includes patients with stable or declining anti-Tg antibody levels without definitive structural evidence of disease.

The majority of published studies examining response to therapy in DTC have been performed in populations of patients whose primary treatment consisted of total thyroidec- tomy and RAI ablation. While the following sections will provide the details of multiple studies examining response to therapy, a simplified overview of the clinical implications of the response-to-therapy reclassification system in patients treated with total thyroidec- tomy and RAI ablation, based on the research from Tuttle et al. (538,539) and supported by other studies, is described in Table 13. Specific suggestions with regard to the application of response-to-therapy assessments in patients treated with total thyroidec- tomy without RAI ablation or with thyroid lobectomy have recently been published by Momesso et al. (602).

[B27] Excellent response: no clinical, biochemical, or structural evidence of disease after initial therapy (remission, NED)

These patients have no clinical, biochemical, or structural evidence of disease identified on risk-appropriate follow-up studies (Table 13). If a total thyroidec- tomy and RAI ablation were done, an excellent response was usually defined as a TSH-stimulated Tg of <1 ng/mL in the absence of structural or functional evidence of disease (and in the absence of anti-Tg antibodies) (538,539,542,601). An excellent response to initial therapy is achieved in 86%–91% of ATA low-risk patients, 57%–63% of ATA intermediate-risk patients, and 14%–16% of ATA high-risk patients (538,539,542).

In 20 retrospective studies, the risk of recurrence over 5–10 years of follow-up ranged from 1% to 4% (median 1.8%) in patients who had an excellent response to therapy by 6–18 months after total thyroidec- tomy and RAI remnant ablation (538,539,542,586–601,603).

The potential impact of this reclassification is the most dramatic in the two-thirds of ATA intermediate-risk patients who achieve an excellent response and therefore have their risk of having recurrent/persistent disease decreased from 36%–43% (predicted by initial ATA risk stratification) to 1%–2% (predicted by response-to-therapy reclassification) (538,539). Because of the very low risk of structural disease recurrence in ATA low-risk patients (1%–3%), reclassification based on an excellent response to therapy has less practical implications than in the intermediate- and high-risk patients.

While many of the studies reviewed primarily low-risk DTC patients (586,588,595–598,601), the same low risk of recurrence following achievement of excellent response to therapy was seen in other studies that had substantial numbers of intermediate-risk patients (538,539,542,587,591,592,604). Furthermore, most studies also demonstrate that the few high-risk patients that achieve an excellent response to therapy also have subsequent recurrence rates in the 1%–2% range (542,593,594,600). The one exception is the 14% risk of recurrence seen in the few ATA high-risk patients that achieved remission in a series from Memorial Sloan-Kettering Cancer Center (538), which is likely due to the known referral bias of unusual and very high-risk cases. Nonetheless, high-risk patients that achieve an excellent response to therapy may require somewhat more intense follow-up than ATA low- and intermediate-risk patients demonstrating an excellent response to therapy. It is important to note that patients at intermediate to high risk of recurrence may require additional structural or functional imaging to rule out disease that may not be detected by US and Tg measurements prior to being classified as having an excellent response (602). The details for choice of follow-up tests are found in another section of these guidelines [C4–C13], (Figs. 5–8).
While most studies have assessed response to therapy using TSH-stimulated Tg values obtained 6–18 months after initial therapy (538,539,542,586,588–590,593–595,597–600), at least 15 other studies have evaluated the response to surgical therapy using a stimulated Tg obtained at the time of RAI remnant ablation (605). Patients demonstrating an excellent response to therapy at this very early time point have a very low risk of disease recurrence (605).

Four additional studies used nonstimulated sensitive Tg assays to define an excellent response to therapy at various time points after initial therapy with total thyroidectomy and RAI remnant ablation (587,595,601,606). A recurrence rate of 1.5% was seen in a cohort of 589 DTC patients who had a nonstimulated Tg <0.27 ng/mL at 3 months after initial therapy (595). Malandrino et al. (587) reported recurrence rates of 0% in low-risk patients, 1% in intermediate-risk patients, and 2.7% in high-risk patients who demonstrated nonstimulated Tg values <0.15 ng/mL at 9–18 months after initial therapy. Smallridge et al. (606) described a 4.3% recurrence rate in 163 low- to intermediate-risk DTC patients with nonstimulated Tg <0.1 ng/mL, measured a median of 1.8 years after initial surgery. Finally, Giovanela et al. (601) reported a 1.6% recurrence rate over 5–6 years of follow-up in 185 low-risk patients who had a nonstimulated Tg of <0.2 ng/mL and normal postoperative neck US 6 months after remnant ablation.

Further studies are needed to refine the precise Tg value cutoff used to define what should be an excellent response to therapy in patients treated with total thyroidectomy with or without RAI remnant ablation. In addition to determining whether TSH-stimulated Tg values are clinically helpful in

<table>
<thead>
<tr>
<th>Table 13. Clinical Implications of Response to Therapy Reclassification in Patients with Differentiated Thyroid Cancer Treated with Total Thyroidectomy and Radioiodine Remnant Ablation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Category</strong></td>
</tr>
<tr>
<td>Excellent response</td>
</tr>
<tr>
<td>Biochemical incomplete response</td>
</tr>
<tr>
<td>Structural incomplete response</td>
</tr>
<tr>
<td>Indeterminate response</td>
</tr>
</tbody>
</table>

NED denotes a patient as having no evidence of disease at final follow-up.
<sup>a</sup>References (538,539).
<sup>b</sup>In the absence of anti-Tg antibodies.
<sup>c</sup>References (583,539,542,586–593,595–601,1078).
<sup>d</sup>References (598,599,617–621).
<sup>e</sup>References (328,607,626,627,898).
FIG. 5. Clinical decision-making and management recommendations in ATA low-risk DTC patients that have undergone total thyroidectomy. R, recommendation in text.

FIG. 6. Clinical decision-making and management recommendations in ATA low-risk DTC patients that have undergone less than total thyroidectomy (lobectomy or lobectomy with isthmusectomy). R, recommendation in text.
patients at very low risk for recurrence, Tg cut points need to be defined for patients whose primary treatment consisted of thyroid lobectomy or total thyroidectomy without RAI ablation.

In summary, once a patient achieves an excellent response to therapy, the initial risk of recurrence estimate should be modified and the patient reclassified as having a subsequent very low risk of recurrence. This reclassification into a very low risk of recurrence status can occur as early as several weeks (or several months) after initial therapy and is applicable to all DTC patients initially stratified as ATA low or intermediate risk of recurrence and to the few ATA high-risk patients that achieve an excellent response to therapy. Appropriate reclassification into the excellent response category with its very low risk of recurrence should lead to re-evaluation of intensity of diagnostic surveillance procedures and treatment, as discussed in other sections of these clinical practice guidelines [C4–C13] (Figs. 5–8).

**B28** Biochemical incomplete response: abnormal Tg values in the absence of localizable disease

These patients have persistently abnormal suppressed and/or stimulated Tg values or rising anti-Tg antibodies without structural evidence of disease that can be detected using risk-appropriate structural and functional imaging (Table 13). Please see section [C6] for discussion of Tg measurements. Previous studies have used nonstimulated Tg values of >1 ng/mL or TSH-stimulated Tg values of >10 ng/mL to define a biochemical incomplete response to therapy in patients treated with total thyroidectomy and RAI ablation (538,539,542). These studies used Tg assays with variable functional sensitivities, so this definition may change over time, especially for the nonstimulated Tg values.

A biochemical incomplete response is not an uncommon outcome and is seen in 11%–19% of ATA low-risk patients, 21%–22% of ATA intermediate-risk patients, and 16%–18% of ATA high-risk patients (538,539). Clinical outcomes in these patients are usually very good, with as many as 56%–68% being classified as having NED at final follow-up, while 19%–27% continue to have persistently abnormal Tg values without structural correlate, and only 8%–17% developing structurally identifiable disease over 5–10 years follow-up (538,539,607). No deaths have been reported in patients with a biochemical incomplete response to therapy followed for up to 10 years (539,607).
Anti-Tg antibody levels measured over time in the same assay can provide clinically useful information (608). Rising anti-Tg antibody titers (or new appearance of anti-Tg antibodies) are associated with an increased risk of disease recurrence (609–614). Conversely, patients rendered free of disease with initial therapy will usually demonstrate a decline in anti-Tg antibody titers over several years (611,615,616).

Vaisman et al. (607) further demonstrated that the transition to NED status occurred without any additional RAI or surgical therapy (beyond LT4 suppressive therapy) in 34% of patients classified as having a biochemical incomplete response to initial therapy. These observations are consistent with several previous studies that have demonstrated that abnormal serum Tg values can gradually decline over time without additional RAI or surgical therapy, in the absence of structurally identifiable disease (598,599,617,617–621).

A small percentage of patients with a biochemical incomplete response to therapy will demonstrate progressive increases in the nonstimulated Tg values over time. In patients treated with total thyroidectomy and RAI remnant ablation, clinically significant increases in unstimulated serum Tg values over time as described by Tg doubling times (<1 year, 1–3 years, or >3 years) (622) or rate of rise in unstimulated Tg of ≥0.3 ng/mL/year over time (623), identify patients at increased risk of developing structurally identifiable loco-regional or distant metastases.

Just as with the excellent response to therapy category, additional studies are needed to more precisely define what Tg levels define an incomplete biochemical response to therapy. The specific cut point that defines an “abnormal Tg” is dependent on the corresponding TSH value, the amount of residual normal thyroid tissue after thyroidectomy, whether or not RAI ablation was performed, and the duration of time since ablation, because Tg values often decline for months to years after ablation. To define a biochemical incomplete response, previous studies have used nonstimulated Tg values of >5 ng/mL at 6 months (624), nonstimulated Tg values >1 ng/mL more than 12 months after ablation (538,539,542), or TSH-stimulated Tg values >10 ng/mL more than 1 year after ablation (538,539). The precise Tg value for defining a biochemical incomplete response to therapy in patients treated with lobectomy or total thyroidectomy without ablation has not been adequately defined. In addition, some studies also classified patients with persistent or rising anti-Tg antibodies in the absence of structurally identifiable disease as having a biochemical incomplete response to therapy (538,539,625).

In summary, a biochemical incomplete response is seen in approximately 15%–20% of DTC patients. Fortunately, many of these patients are eventually reclassified as having NED at final follow-up, often without any additional RAI or surgical treatments.

[B29] Structural incomplete response: persistent or newly identified loco-regional or distant metastases

These patients have structural or functional (RAI scan, 18F-FDG-PET) evidence of loco-regional or distant metastases.
(538,539,607). This category includes both patients with biopsy-proven disease and also patients in whom structural or functional disease is identified, which is highly likely to be metastatic disease based on the clinical scenario (Table 13). A structural incomplete response to initial therapy is seen in 2%–6% of ATA low-risk patients, 19%–28% of ATA intermediate-risk patients, and 67%–75% of ATA high-risk patients (538,539).

Despite additional treatments, the majority of patients classified as having a structural incomplete response will have persistent structural and/or biochemical evidence of persistent disease at final follow-up (539,607). Depending on the definition used to describe patients as free from disease, higher rates of remission (29%–51%) have been described following surgical intervention for patients with persistent/recurrent loco-regional disease (626–628). While no deaths were reported over a follow-up period that extended to 15 years in patients with biochemical incomplete response to therapy, death from disease was seen in 11% of patients with a loco-regional incomplete response and in 57% of patients with structurally identifiable distant metastases (539,607).

In summary, a structural incomplete response to initial therapy identifies a cohort of DTC patients that may not be cured with additional therapies and consequently demonstrate the highest risk of disease-specific mortality of any of the response-to-therapy categories. Persistent/recurrent loco-regional structural disease may have a higher likelihood of responding to additional treatments and has significantly lower disease-specific mortality rates than persistent/recurrent distant metastases.

**[B30] Indeterminate response: biochemical or structural findings that cannot be classified as either benign or malignant (acceptable response)**

Patients with an indeterminate response have biochemical, structural, or functional findings that cannot be confidently classified as either excellent response or persistent disease (328,538,539,606,629) (Table 13). Rather than forcing these patients into either the excellent or incomplete response-to-therapy categories, some investigators have recommended a separate category for these patients so that they can be continued to be carefully observed, with selected patients identified for further evaluation with testing designed to establish the presence or absence of disease (538,539).

For example, this category includes patients with subcentimeter avascular thyroid bed nodules or atypical cervical lymph nodes that have not been biopsied, faint uptake in the thyroid bed with undetectable stimulated Tg on follow-up imaging, or nonspecific abnormalities on functional or cross-sectional imaging. Also included in this category are patients with nonstimulated Tg values that are detectable but <1 ng/mL, TSH-stimulated Tg values between 1 and 10 ng/mL, and stable or declining Tg antibodies in the same assay over time in the absence of structural disease (538,539).

This issue was exemplified in a recent study evaluating the prognostic value of a highly sensitive Tg assay in which the response to therapy could not be definitively established in 16 patients that had small indeterminate pulmonary micronodules without other evidence for persistent disease (606). A similar situation arises when trying to determine the response to therapy in the 34% of patients that demonstrated nonspecific subcentimeter thyroid bed nodules after total thyroidectomy (629). Similarly, it is often difficult to be certain whether or not very low-level detectable Tg values represent persistent disease or simply remnant normal thyroid cells remaining after initial therapy.

An indeterminate response to initial therapy is seen in 12%–29% of ATA low-risk patients, 8%–23% of ATA intermediate-risk patients, and 0%–4% of ATA high-risk patients (538,539). The clinical outcomes in patients with an indeterminate response to therapy are intermediate between patients with an excellent response and those with incomplete responses. Two series have demonstrated that only 13%–20% of patients with an indeterminate response to therapy are reclassified as persistent/recurrent disease over approximately 10 years of follow-up. In the remaining 80%–90% of patients, the nonspecific findings either remain stable or resolve with observation alone.

In summary, the majority of patients with an indeterminate response to therapy remain disease-free during prolonged follow-up. However, up to 20% of these patients will eventually have biochemical, functional, or structural evidence of disease progression and may require additional therapies.

**[B31] Using risk stratification to guide disease surveillance and therapeutic management decisions**

Risk stratification is the cornerstone of individualized thyroid cancer management. Initial risk estimates are useful to guide the wide variety of clinical management decisions that need to be made around the time of initial diagnosis and treatment. As described in this document, initial management decisions are largely made by balancing the estimates of the risk of recurrence and the risk of disease-specific mortality with the potential benefits and risks of proposed therapies. However, in clinical practice many other risk estimates can also significantly influence surveillance and therapeutic decision-making, including the risk of failing initial therapy, the risk of having non–RAI-avid disease, the risk of disease recurring without making appreciable amounts of serum Tg, the risk of adjuvant RAI therapy, the risk of additional thyroid surgery, the risk of additional lymph node surgery, the risk of external beam radiation therapy, and the risk of systemic therapy. Individual management recommendations require that the risks and benefits of potential surveillance and therapeutic management decisions be carefully evaluated in the context of the specific clinicopathologic features of each patient.

Nonetheless, initial risk estimates can be used to guide recommendations with regard to the extent of thyroid surgery, the need for and extent of cervical lymph node dissection, the need for and the dose of administered activities of RAI, the need for and degree of TSH suppression, the need for and details of external beam radiation therapy, the need for and types of systemic therapy, the need for and types of studies required for initial staging, and the intensity and type of follow-up studies required for evaluating response to therapy in the early years following initial therapy. This approach tailors the aggressiveness of intervention and follow-up to the specific risks associated with the tumor in an individual patient.

In summary, this risk-adapted management approach utilizes initial risk estimates to guide early surveillance and therapeutic management decisions. These initial management plans are then modified over time as additional data accumulate and allow for restratification based on individual response to therapy. This system tailors the extent and
intensity of therapy and follow-up studies to real-time risk estimates that evolve over time for individual patients.

[B32] Should postoperative disease status be considered in decision-making for RAI therapy for patients with DTC?

**RECOMMENDATION 50**

(A) Postoperative disease status (i.e., the presence or absence of persistent disease) should be considered in deciding whether additional treatment (e.g., RAI, surgery, or other treatment) may be needed.

(Strong recommendation, Low-quality evidence)

(B) Postoperative serum Tg (on thyroid hormone therapy or after TSH stimulation) can help in assessing the persistence of disease or thyroid remnant and predicting potential future disease recurrence. The Tg should reach its nadir by 3–4 weeks postoperatively in most patients.

(Strong recommendation, Moderate-quality evidence)

(C) The optimal cutoff value for postoperative serum Tg or state in which it is measured (on thyroid hormone therapy or after TSH stimulation) to guide decision-making regarding RAI administration is not known.

(No recommendation, Insufficient evidence)

(D) Postoperative diagnostic RAI WBSs may be useful when the extent of the thyroid remnant or residual disease cannot be accurately ascertained from the surgical report or neck ultrasonography, and when the results may alter the decision to treat or the activity of RAI that is to be administered. Identification and localization of uptake foci may be enhanced by concomitant single photon emission computed tomography–computed tomography (SPECT/CT). When performed, pretherapy diagnostic scans should utilize $^{123}$I (1.5–3 mCi) or a low activity of $^{131}$I (1–3 mCi), with the therapeutic activity optimally administered within 72 hours of the diagnostic activity.

(Weak recommendation, Low-quality evidence)

Postoperative disease status is a relevant consideration in postoperative treatment decision-making after initial consideration of clinic-pathologic stage. Evaluation of postoperative disease status may be performed by a number of means including serum Tg, neck ultrasonography, and iodine radioisotope scanning. There are currently no RCTs comparing any particular postoperative diagnostic strategy with the intention of modulating decision-making on RAI remnant ablation or RAI treatment for DTC.

[B33] Utility of postoperative serum Tg in clinical decision-making

Serum Tg measurements (with anti-Tg antibodies), with or without neck US, are frequently performed as part of the early postoperative evaluation. Please see section [C6] for discussion of Tg measurements. The predictive value of the postoperative Tg value will be significantly influenced by a wide variety of factors including the amount of residual thyroid cancer and/or normal thyroid tissue, the TSH level at the time of Tg measurement, the functional sensitivity of the Tg assay, the Tg cutoff used for analysis, the individual risk of having RAI-avid loco-regional or distant metastasis, the time elapsed since total thyroidectomy, and/or the sensitivity of the posttherapy scanning technique (SPECT/CT vs. planar imaging).

Multiple studies have confirmed an increase risk of recurrence following total thyroidectomy and RAI remnant ablation in patients that had a postoperative TSH-stimulated Tg $>1–2$ ng/mL at the time of ablation (596,605,630–637). In multivariate analysis, the postoperative Tg is often found to be an independent predictor of persistent or recurrent disease (596,630,631,636,637). Furthermore, the risk of having recurrent or persistent disease increases as the postoperative Tg rises (634,636). Using receiver operator curve analyses, thyroid hormone withdrawal postoperative Tg values between 20 and 30 ng/mL achieve the optimal balance of sensitivity and specificity for predicting recurrent or persistent disease (638–640). Furthermore, high postoperative stimulated Tg values ($>10–30$ ng/mL) are also associated with poorer survival (636,639,641). Conversely, postoperative stimulated Tg values less than $1–2$ ng/mL are strong predictors of remission (634,636). Even in ATA low- and intermediate-risk patients that did not receive RAI remnant ablation, a nonstimulated postoperative Tg $<1$ ng/mL was associated with excellent clinical outcomes and recurrence rates $<1\%$ (642). The median follow-up in this study was 62 months (2–116 months). Therefore, a postoperative serum Tg can provide valuable information with regard to the likelihood of achieving remission or having persistent or recurrent disease in response to an initial therapy.

A postoperative Tg $<10$ ng/mL may not distinguish between nodal disease and thyroid remnant, when evaluated using concurrent RAI scans with SPECT/CT (643). In one of the prospective studies mentioned previously, a postoperative Tg threshold of $>5$ ng/mL was suggested as an indication for RAI treatment (633). However, in a recent retrospective review of consecutive low-risk patients treated with total thyroidectomy without RAI, an unstimulated Tg of $>2$ ng/mL with a concomitant median TSH level of 0.48 mIU/L was reported to detect all patients with disease recurrence (76 patients followed for a median of 2.5 years) (644). Thus, there is some uncertainty as to what degree of postoperative stimulated or unstimulated thyroglobulinemia (with or without neck US interpretation) may be appropriate to prompt RAI treatment. Moreover, detection of unexpected inappropriate thyroglobulinemia may prompt consideration of further investigation for its cause (e.g., imaging studies).

The postoperative Tg can also be used to predict the likelihood of identifying RAI-avid metastatic thyroid cancer outside the thyroid bed on the posttherapy scan at the time of remnant ablation. No uptake outside the thyroid bed was detected in 63 low-risk patients with a nonstimulated postoperative Tg of $<0.4$ ng/mL (630) or in 132 low-risk patients with a thyroid hormone withdrawal Tg of $<1$ ng/mL (645). However, RAI-avid metastatic foci outside the thyroid bed were detected in 12% of intermediate-risk patients with a suppressed Tg of $<0.6$ ng/mL (646), 5.6% of intermediate/high risk patients with a suppressed Tg of $<1.0$ ng/mL (647), and 6.3% of intermediate/high-risk patients with a thyroid hormone withdrawal stimulated Tg of $<2$ ng/mL (648). The likelihood of finding RAI-avid metastatic disease on the posttherapy scan is substantially lower (2.8%) if the postoperative
Tg is undetectable in three different Tg assays than if it is undetectable only in a single assay (30%) (649). Conversely, the likelihood of identifying either loco-regional or distant metastases on the posttherapy scan increases as either the suppressed or stimulated Tg values rise above 5–10 ng/mL (631,646,647,650). Therefore, neither a stimulated or suppressed postoperative Tg of <1 ng/mL can completely eliminate the possibility that a posttherapy RAI scan will identify metastatic foci outside the thyroid bed. However, postoperative Tg values greater than 5–10 ng/mL increase the likelihood of identifying RAI-avid metastatic disease on the posttherapy scan.

The postoperative serum Tg value can also be used to predict the likelihood of successful remnant ablation. Postoperative thyroid hormone withdrawal stimulated Tg values >5–6 ng/mL were associated with higher rates of failed ablation after administered activities of both 30 mCi (651) and 100 mCi (652). A TSH-stimulated Tg >6 ng/mL was associated with a 5-fold greater risk of failing ablation after an activity of 30 mCi administered after preparation with thyroid hormone withdrawal (651).

It does appear that a postoperative Tg value (either TSH-stimulated or nonstimulated) is an important prognostic factor that can be used to guide clinical management. Given a disappearance half-life of 1–3 days (653–658), the postoperative Tg should reach its nadir by 3–4 weeks postoperatively in nearly all patients. In low-risk patients, a suppressed or stimulated Tg <1 ng/mL is very reassuring and further confirms classification of the patients as being at low risk. In intermediate-risk patients, postoperative Tg values <1 ng/mL are reassuring, but do not completely rule out the presence of small-volume RAI-avid metastatic disease. However, even without RAI ablation, many intermediate risk patients have excellent clinical outcomes. Therefore, it is not clear that additional therapy is required in these intermediate-risk patients with postoperative Tg values <1 ng/mL even though small-volume RAI-avid disease may still be present after thyroidectomy.

On the other hand, postoperative Tg values (stimulated or nonstimulated) greater than 10–30 ng/mL increase the likelihood of having persistent or recurrent disease, failing initial RAI ablation, having distant metastases, and dying of thyroid cancer. Therefore, postoperative Tg values >10 ng/mL will likely lead to additional evaluations and possibly even additional therapies.

With regard to decision-making on the need for RAI remnant ablation, it appears that the postoperative serum Tg value will be more helpful in identifying patients that may benefit from RAI ablation rather than in identifying patients that do not require ablation. For example, a postoperative Tg value >5–10 ng/mL may lead to selection of RAI ablation in an ATA low-risk patient or ATA intermediate-risk patient that otherwise would not have required RAI ablation (selective use) in order to improve initial staging and facilitate follow-up. Conversely, in high-risk patients, a postoperative Tg value <1 ng/mL does not rule out RAI-avid disease and therefore is unlikely to alter the decision to proceed with RAI ablation.

[B34] Potential role of postoperative US in conjunction with postoperative serum Tg in clinical decision-making

In a prospective study of 218 DTC patients, Lee et al. (659) reported that a stimulated Tg <2 ng/mL after thyroid hormone withdrawal (with goal TSH of >30 mIU/L), at the time of administration of 100–200 mCi of $^{131}$I (for remnant ablation or treatment), was associated with the following NPVs for biochemical or structural recurrence at 6–12 months: 98.4% for ATA low-risk patients, 94.1% for ATA intermediate-risk patients, and 50% in the ATA high-risk group. They further reported that the NPVs increased to 97.2%, and 100% for ATA intermediate- and high-risk patients, respectively, when the stimulated Tg values were combined with negative neck US findings at baseline (with no change in low-risk patients). Similar significant decreases in the risk of recurrence were seen when ATA intermediate- and high-risk patients had a normal postoperative neck US (660).

[B35] Role of postoperative radioisotope diagnostic scanning in clinical decision-making

Iodine radioisotope diagnostic testing may include $^{131}$I or $^{123}$I diagnostic imaging with or without SPECT-CT, and/or RAI uptake measurements. Postoperative RAI planar imaging (123I or 131I, with or without SPECT-CT) has been reported to yield information that could alter clinical management (such as altering disease status assessment) in 25%–53% of patients, as reported in single-center, retrospective studies (643,661,662). However, in a multivariate analysis of retrospective data, Hu et al. (663) reported that the use of 5 mCi of 131I between 4 and 11 days prior to remnant ablation was independently associated with an increased risk of remnant ablation failure. In contrast, in a smaller retrospective study, the administration of 3–5 mCi of 131I for scanning 2–5 days prior to ablation in 37 patients was not associated with any significant reduction in remnant ablation success, compared to no pretherapy scanning in 63 patients (131I therapeutic activity of 100–200 mCi used in both groups) (664). A possible relationship between 131I diagnostic scan activity on remnant ablation success was suggested in another retrospective study, in which success was lower following the use of 3 mCi as compared to 1 mCi of 131I, 9 days before therapeutic administration of 100 mCi (665). In two small RCTs, there was no significant impact of 131I scanning compared to 123I scanning on the rate of successful remnant ablation (666,667). The timing of whole-body diagnostic scans following administration of radioisotopes in reviewed studies ranged from about 24 to 72 hours for 131I (643,662,663,665–667) and was 24 hours for 123I (661,662,666). The tailoring of RAI therapeutic activity according to RAI neck uptake (measured 24 hours after administration of 1 mCi of 131I) was associated with a lower rate of remnant ablation success than fixed dosing, in another single-center retrospective observational study (668). Furthermore, in a multivariate analysis of retrospective data (adjusted for relevant risk factors), Verburg et al. (669) reported that the use of 1 mCi of 131I for calculation of RAI neck uptake 2 days before remnant ablation was independently associated with an increased risk of remnant ablation failure, although Yap and Murby showed that 1.1 mCi 131I diagnostic scans did not adversely affect the success of ablation or recurrence rate at 3 years (670).

There continues to be discussion on the utility of postoperative iodine radioisotope diagnostic scanning (with or without SPECT/CT) in guiding RAI therapeutic decision-making. Valuable information on disease status, remnant uptake, and the presence of residual RAI-avid disease may be obtained by such testing, which could alter management and potentially
benefit outcome. Questions regarding the potentially negative impact of such scans with $^{131}$I on RAI therapeutic efficacy for successful remnant ablation (“stunning”) may be mitigated or avoided by the use of either low-activity $^{131}$I (1–3 mCi) or alternative isotopes such as $^{123}$I.

[B36] **What is the role of RAI (including remnant ablation, adjuvant therapy, or therapy for persistent disease) after thyroidectomy in the primary management of DTC?**

### RECOMMENDATION 51 (details in Table 14)

(A) RAI remnant ablation is not routinely recommended after thyroidectomy for ATA low-risk DTC patients. Consideration of specific features of the individual patient that could modulate recurrence risk, disease follow-up implications, and patient preferences are relevant to RAI decision-making.

(Weak recommendation, Low-quality evidence)

(B) RAI remnant ablation is not routinely recommended after lobectomy or total thyroidectomy for patients with unifocal papillary microcarcinoma, in the absence of other adverse features.

(Strong recommendation, Moderate-quality evidence)

(C) RAI remnant ablation is not routinely recommended after thyroidectomy for patients with multifocal papillary microcarcinoma in absence of other adverse features. Consideration of specific features of the individual patient that could modulate recurrence risk, disease follow-up implications, and patient preferences are relevant to RAI decision-making.

(Weak recommendation, Low-quality evidence)

(D) RAI adjuvant therapy should be considered after total thyroidectomy in ATA intermediate-risk level DTC patients.

(Weak recommendation, Low-quality evidence)

(E) RAI adjuvant therapy is routinely recommended after total thyroidectomy for ATA high risk DTC patients.

(Strong recommendation, Moderate-quality evidence)

Depending on the postoperative risk stratification of the individual patient, the primary goal of postoperative administration of RAI after total thyroidectomy may include (i) RAI remnant ablation (to facilitate detection of recurrent disease and initial staging by tests such as Tg measurements or whole-body RAI scans), (ii) RAI adjuvant therapy (intended to improve disease-free survival by theoretically destroying suspected, but unproven residual disease, especially in patients at increased risk of disease recurrence), or (iii) RAI therapy (intended to improve disease-specific and disease-free survival by treating persistent disease in higher risk patients). Additional considerations in RAI decision-making may include patient comorbidities (and the potential impact of therapeutic doses of RAI or preparation for the procedure), feasible or preferred disease surveillance procedures, patient preferences (the latter being particularly important when clear data on therapeutic efficacy are lacking), or others. It is important to note that in patients with low-risk DTC, disease surveillance may be accomplished without RAI ablation using neck US and Tg with Tg antibody measurements while on thyroid hormone therapy.

We categorized the results of our review according to the ATA Risk of Recurrence Risk stratification (outlined in a preceding section of these guidelines). However, given that the ATA risk classification is relatively new and the majority of studies examining therapeutic efficacy of postsurgical RAI remnant ablation or therapy (adjuvant or for persistent disease) have been performed with attention to traditional mortality risk stratification systems such as the AJCC/TNM system, MACIS, National Thyroid Cancer Treatment Cooperative Study Group (NTCTCSG), or others, it was necessary to extrapolate the results of many studies according to estimated ATA risk level. We have also categorized some of the results of our evidence review according to the AJCC/TNM risk of mortality stratification system because this system has been in use longer in our field (Table 14). Evaluation of postoperative disease status and recommendations for RAI remnant ablation and adjuvant therapy can be found in algorithms in Figs. 5–8.

**ATA low risk.** Studies examining the impact of RAI remnant ablation/adjuvant therapy on long-term thyroid cancer outcomes in ATA low-risk patients are subject to limitations due to their observational nature (and potential for bias), as well as limited statistical power to detect relatively uncommon events (such as disease-related mortality). By definition, the risk of disease-specific mortality is low, the risk of persistent/recurrent disease is low (around 3%), and there is no evidence that delayed discovery and treatment of persistent disease may decrease the chance of cure in these patients. In a retrospective multicenter registry study, 1298 DTC patients categorized as being in the ATA low-risk level, were followed for a median of 10.3 years, and there was no significant effect of RAI adjuvant therapy on overall or disease-free survival, using respective multivariate and stratified propensity analysis techniques (544). Prospective data from the NTCTCSG suggest that overall disease-specific and disease-free survival are not improved by RAI treatment in NTCTCSG stage I and II patients (i.e., patients aged <45 years with no distant metastases or patients aged ≥45 years with a primary tumor <4 cm in diameter, no extrathyroidal extension, and no nodal metastases), also using multivariate analyses and propensity analyses (671,672). In two systematic reviews examining results of multivariate adjusted analyses, with a focus on low-risk DTC using classic clinicopathologic staging systems such as TNM/AJCC, the majority of studies did not show a significant effect of RAI adjuvant therapy in reducing thyroid cancer–related death, and conflicting findings of studies relating to outcomes of disease recurrence (673,674). A more recent systematic review of the literature supported the findings of the earlier systematic reviews (675). It is important to note that in the studies summarized in this section on ATA low-risk disease (or equivalent), patients with multifocal PTC were generally included (if no other adverse features meeting criteria for upstaging were noted). To date, there is little evidence to suggest that RAI may improve disease-specific mortality in low-risk DTC patients, and there is some conflicting evidence.
**Table 14. Characteristics According to the American Thyroid Association Risk Stratification System**

<table>
<thead>
<tr>
<th>ATA risk Staging (TNM)</th>
<th>Description</th>
<th>Body of evidence suggests RAI improves disease-specific survival?</th>
<th>Body of evidence suggests RAI improves disease-free survival?</th>
<th>Postsurgical RAI indicated?</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATA low risk T1a N0,Nx M0,Mx</td>
<td>Tumor size ≤1 cm (uni-or multi-focal)</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>ATA low risk T1b,T2 N0, Nx M0,Mx</td>
<td>Tumor size &gt;1–4 cm</td>
<td>No</td>
<td>Conflicting observational data</td>
<td>Not routine&lt;sup&gt;b&lt;/sup&gt;—May be considered for patients with aggressive histology or vascular invasion (ATA intermediate risk).</td>
</tr>
<tr>
<td>ATA low to intermediate risk T3 N0,Nx M0,Mx</td>
<td>Tumor size &gt;4 cm</td>
<td>Conflicting data</td>
<td>Conflicting observational data</td>
<td>Consider&lt;sup&gt;b&lt;/sup&gt;—Need to consider presence of other adverse features. Advancing age may favor RAI use in some cases, but specific age and tumor size cutoffs subject to some uncertainty.&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>ATA low to intermediate risk T3 N0,Nx M0,Mx</td>
<td>Microscopic ETE, any tumor size</td>
<td>No</td>
<td>Conflicting observational data</td>
<td>Consider&lt;sup&gt;b&lt;/sup&gt;—Generally favored based on risk of recurrent disease. Smaller tumors with microscopic ETE may not require RAI.</td>
</tr>
<tr>
<td>ATA low to intermediate risk T1-3 N1a M0,Mx</td>
<td>Central compartment neck lymph node metastases</td>
<td>No, except possibly in subgroup of patients ≥45 years of age (NTCTCSG Stage III)</td>
<td>Conflicting observational data</td>
<td>Consider&lt;sup&gt;b&lt;/sup&gt;—Generally favored, due to somewhat higher risk of persistent or recurrent disease, especially with increasing number of large (&gt;2–3 cm) or clinically evident lymph nodes or presence of extranodal extension. Advancing age may also favor RAI use.&lt;sup&gt;a&lt;/sup&gt; However, there is insufficient data to mandate RAI use in patients with few (&lt;5) microscopic nodal metastases in central compartment in absence of other adverse features.</td>
</tr>
<tr>
<td>ATA low to intermediate risk T1-3 N1b M0,Mx</td>
<td>Lateral neck or mediastinal lymph node metastases</td>
<td>No, except possibly in subgroup of patients ≥45 years of age</td>
<td>Conflicting observational data</td>
<td>Consider&lt;sup&gt;b&lt;/sup&gt;—Generally favored, due to higher risk of persistent or recurrent disease, especially with increasing number of macroscopic or clinically evident lymph nodes or presence of extranodal extension. Advancing age may also favor RAI use.&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>ATA high risk T4 Any N Any M</td>
<td>Any size, gross ETE</td>
<td>Yes, observational data</td>
<td>Yes, observational data</td>
<td>Yes</td>
</tr>
<tr>
<td>ATA high risk M1 Any T Any N</td>
<td>Distant metastases</td>
<td>Yes, observational data</td>
<td>Yes, observational data</td>
<td>Yes</td>
</tr>
</tbody>
</table>

<sup>a</sup>Recent data from the NTCTCSG (National Thyroid Cancer Treatment Cooperative Study Group) have suggested that a more appropriate prognostic age cutoff for their and other classification systems could be 55 years, rather than 45 years, particularly for women.

<sup>b</sup>In addition to standard clinicopathologic features, local factors such as the quality of preoperative and postoperative US evaluations, availability and quality of Tg measurements, experience of the operating surgeon, and clinical concerns of the local disease management team may also be considerations in postoperative RAI decision-making.
on effect on recurrence, with newer data using the ATA risk system suggesting the lack of a significant effect. Furthermore, the majority of the best available observational evidence suggests that RAI adjuvant therapy is unlikely to improve disease-specific or disease-free survival in papillary microcarcinoma (<1 cm, uni- or multifocal), in the absence of other higher risk features (146,676–680). In a recent retrospective analysis of 704 papillary microcarcinoma patients whose initial risk level was ATA low or intermediate who were followed for a median of 64 months, there was no significant reduction in recurrence rates in patients treated with RAI compared to those not treated with RAI using a propensity score analysis (677). With respect to microcarcinomas of follicular cancer and Hürthle cell cancer, a recent SEER registry secondary data analysis suggested no disease-specific survival benefit in patients treated with RAI in a multivariate analysis adjusted for age, histology, disease extent, type of surgery, and external beam radiation therapy (the total number of patients in this study was 564) (681). A limitation of interpreting these data on follicular and Hürthle cell microcarcinomas is that some of the patients in the study had some adverse features and were not all considered low risk; however, the authors adjusted for relevant variables in their multivariate analysis (681). The role of RAI adjuvant therapy in ATA low-risk DTC should be clarified in the future, following completion of RCTs, such as the Iodine or Not (IoN) trial for low- and intermediate-risk patients (682) and ESTIMABL2 for low-risk patients.

**ATA intermediate risk.** Multivariate adjusted analyses from SEER suggest that postsurgical RAI treatment is associated with improved overall survival for aggressive PTC histologies such as tall cell, diffuse sclerosing, and insular variants (683,684). Furthermore, multivariate adjusted analyses from SEER suggest that RAI treatment is associated with improved overall survival in node-positive adult patients with PTC or pT3 node-negative PTC, in which the primary tumor is >4 cm or there is evidence of microscopic extrathyroidal extension (685). It is important to note, however, that in this SEER study the overall survival rate was very high in node-positive or pT3 node-negative PTC patients aged <45 years, such that 99% and 98% of such individuals were alive after a median follow-up period of 6.8 years, with or without RAI treatment, respectively. The clinical significance of this approximately 1% absolute risk difference could be questioned. In contrast, in the same study, for individuals aged ≥65 years, assuming the same median study follow-up period of 6.8 years, 73% of T3-node negative or node-positive PTC patients treated with RAI and 69% of those not treated with RAI were alive (685); in this older subgroup, the absolute risk difference would be estimated to be about 4%.

There is some supportive evidence from multivariate and propensity analyses that there may be a benefit of adjuvant RAI treatment in improving overall and disease-specific survival as well as disease-free survival in patients with nodal metastases aged ≥45 years, as such patients would be included in the NTCTCSG stage III category (671). Furthermore, in a single-center retrospective study from Hong Kong examining data from a subgroup of 421 patients with node-positive PTC, lymph node failure-free survival was improved with postsurgical RAI treatment, with the greatest treatment benefits observed in patients with N1b disease, as well as with lymph nodes >1 cm in diameter (686). RAI therapeutic efficacy in patients <45 years of age with nodal metastases is unclear because such patients are categorized as stage I by the NTCTCSG system and no significant benefit of RAI treatment was observed for the stage I group in that study, with no specific subgroup analysis reported according to node positivity (671). A single-center retrospective study from the Mayo Clinic examining 20-year cause-specific mortality and recurrence rate, suggested no significant benefit in PTC patients with a MACIS score of ≤6 who had positive lymph nodes, using respective univariate analyses (278). Given that age is incorporated in the MACIS score, it is possible that age was a contributing factor in the results of that analysis. In a subgroup of 352 patients with microscopic extrathyroidal extension from a single-center retrospective study, postsurgical RAI treatment was associated with a reduction in rate of local relapse (686). However, in the NTCTCSG study, microscopic extraglandular invasion would be classified as NTCTCSG stage I for patients <45 years of age and II for those ≥45 years of age, and those stages were associated with lack of clear benefit of RAI. In a recent systematic review, Lamartina et al. (675) reported conflicting results on the impact of RAI treatment on disease recurrence, specifically indicating that 11 nonrandomized studies suggested a benefit, whereas 13 studies did not show a significant benefit. For patients with ATA intermediate-risk DTC, limited risk-group specific data examining RAI efficacy is available, but existing data suggest that the greatest potential benefit may be observed with adverse thyroid cancer histologies, increasing volume of nodal disease, lymph node disease outside the central neck, and advancing patient age. Benefits regarding survival or recurrence can be expected primarily in patients with higher risk of recurrent or persistent disease that is iodine avid. More studies are needed, including RCTs, to characterize RAI treatment efficacy in ATA intermediate-risk patients. The adjuvant therapeutic efficacy of RAI treatment in improving long-term thyroid cancer outcomes in the situation of isolated microscopic central neck nodal disease in the absence of other adverse features is unknown, so the relatively good overall prognosis of this group (as outlined in the preceding section of these guidelines) as well as the uncertain RAI therapeutic efficacy for this subgroup, are important considerations in RAI decision-making. Clearly more research is needed to understand the therapeutic efficacy in various subgroups of patients in the ATA intermediate-risk category.

**ATA high risk.** A prospective multicenter study reported a significant improvement in overall and disease-specific mortality, as well as disease-free survival in NTCTCSG stage III and IV patients, after statistical adjustment using multivariate and propensity stratified analyses (671). Furthermore, prospectively collected data from the SEER cancer registry suggest that postsurgical RAI therapy is associated with improved overall survival in patients with PTC with distant metastases (when distant metastases were combined with age >45 years, tumor size >2 cm, and positive lymph nodes at primary diagnosis) (687). Data from SEER also suggest that overall survival in patients with FTC with distant metastases more than doubled in patients receiving postsurgical RAI treatment (687). Thus, routine postsurgical RAI treatment is recommended in patients with ATA high-risk DTC.
RECOMMENDATION 52

The role of molecular testing in guiding postoperative RAI use has yet to be established; therefore, no molecular testing to guide postoperative RAI use can be recommended at this time.

(No recommendation, Insufficient evidence)

Preclinical studies show that the presence of the BrafV600E mutation significantly reduces sodium-iodide symporter expression and RAI uptake (688). There are currently insufficient clinical data, however, to know whether the presence or absence of the BrafV600E mutation or other genetic alterations in PTC may impact the success of adjuvant therapy or remnant ablation in PTC, or if adjustments in administered activity are warranted for any planned treatments. In a recent subgroup analysis of 134 PTC patients with T1aN0M0 disease, the rate of macroscopic structural disease recurrence was 0% in the no RAI group (24% of whom were Braf positive), 2.6% in 39 Braf-positive patients who received RAI, and 1.7% in Braf-negative patients who received RAI; the rate of biochemical persistence of disease (defined by a stimulated Tg of >1 ng/mL) was 13% in the 39 Braf-positive patients and 1.7% in the Braf-negative patients; the standard activity for remnant ablation in this study was 30 mCi in most cases (565). The relatively small number of patients who did not receive postoperative RAI and the relatively small number of structural disease recurrences in the T1aN0M0 subgroup of PTC patients and lack of randomization in this study may preclude meaningful analysis of RAI therapeutic efficacy. The ESTIMABL2 study will analyze the relevance of Braf status on outcome (registration number NCT01837745). There are no studies examining therapeutic efficacy of RAI in ATA high-risk patients, but the presence or absence of a BrafV600E mutation in that situation would be unlikely to alter RAI decision-making at present. Thus, the role of molecular testing in guiding postoperative RAI use has yet to be established, and more research in this area is clearly needed. Moreover, in general, RCTs examining RAI therapeutic efficacy are needed, and ideally these should be appropriately stratified for ATA recurrence risk level and other important prognostic variables.

B37] What is the role of molecular marker status in therapeutic RAI decision-making?

RECOMMENDATION 53

(A) If thyroid hormone withdrawal is planned prior to RAI therapy or diagnostic testing, LT4 should be withdrawn for 3–4 weeks. Liothyronine (LT3) may be substituted for LT4 in the initial weeks if LT4 is withdrawn for 4 or more weeks, and in such circumstances, LT3 should be withdrawn for at least 2 weeks. Serum TSH should be measured prior to radioisotope administration to evaluate the degree of TSH elevation.

B38] How long does thyroid hormone need to be withdrawn in preparation for RAI remnant ablation/treatment or diagnostic scanning?

(Strong recommendation, Moderate-quality evidence)

(B) A goal TSH of >30 mIU/L has been generally adopted in preparation for RAI therapy or diagnostic testing, but there is uncertainty relating to the optimum TSH level associated with improvement in long-term outcomes.

(Weak recommendation, Low-quality evidence)

Thyrotropin stimulation before RAI remnant ablation/therapy or scanning has been a long-established standard of care because early observational research suggested that a TSH >30 mIU/L was required for incompletely resected thyroid tumors to significantly concentrate 131I (689). There have been two RCTs comparing various thyroid hormone withdrawal protocols prior to therapeutic or diagnostic iodine radioisotope administration (690,691). Lee et al. (691) reported on an open-label, single-center study, in which 291 patients with well-differentiated thyroid cancer (TNM stage T1–T3, N0/N1a,M0) were randomized to either (a) withdrawal LT4 for 4 weeks (n = 89), (b) withdrawal of LT4 for 4 weeks with substitution of LT3 for the first 2 weeks (n = 133), or (c) recombinant human TSH (rhTSH; with withdrawal of LT4 for a few days from the time of the first rhTSH injection to radioisotope administration) (n = 69) (691). In this trial, all patients received 30 mCi of 131I for remnant ablation and were prescribed a 2-week low-iodine diet (LID) pre-ablation. Although the randomization method was unclear, the baseline characteristics (including pre-ablation urinary iodine measurements) were well balanced among groups. Furthermore, the pre-ablation TSH was >30 in all patients in all groups in this trial, with no significant difference in mean pre-ablation TSH levels. Moreover, the primary outcome, which was the rate of successful remnant ablation at 12 months, was not significantly different among groups (range 91.0%–91.7% among groups). Upon administration of questionnaires in a double-blind fashion, there was no significant difference in quality of life during preparation for RAI ablation, between the LT4 withdrawal group and the LT3 withdrawal with LT3 substitution group; however, quality of life in both withdrawal groups prior to remnant ablation was significantly worse than after rhTSH preparation (691). Long-term outcome data from this trial were not reported. In a single-center trial, Leboeuf et al. (690) randomized 20 individuals with well-differentiated thyroid cancer awaiting RAI remnant ablation or diagnostic scanning to LT3 withdrawal and either (a) substitution of LT3 (50 μg/d, divided as two capsules) for 21 days, followed by 2 weeks off LT3, or (b) identical-appearing placebo for LT3 (two pills per day) for 21 days. In both groups, either the LT3 or placebo was withdrawn for another 2 weeks, and weekly measurements were performed for serum TSH, free thyroxine, and free triiodothyronine (690). The primary outcome was the hypothyroidism symptom score (Billewicz scale), which was ascertained in a double-blind fashion at time of LT4 withdrawal and every 2 weeks until the end of the study. The randomization method was a computer-generated number sequence; the LT3 group was significantly older than the placebo group (mean age 64 compared to 46), suggesting imbalance in the randomization (690). Disease stage of participants was not reported. Approximately 15% of participants withdrew from this trial (two in the placebo group and one in the LT3 group). Leboeuf et al. (690) reported no significant differences between the two thyroid hormone
RECOMMENDATION 54

(A) In patients with ATA low-risk and ATA intermediate-risk DTC without extensive lymph node involvement (i.e., T1–T3, N0/Nx/N1a, M0), in whom RAI remnant ablation or adjuvant therapy is planned, preparation with rhTSH stimulation is an acceptable alternative to thyroid hormone withdrawal for achieving remnant ablation, based on evidence of superior short-term quality of life, noninferiority of remnant ablation efficacy, and multiple consistent observations suggesting no significant difference in long-term outcomes.

(Strong recommendation, Moderate-quality evidence)

(B) In patients with ATA intermediate-risk DTC who have extensive lymph node disease (multiple clinically involved LN) in the absence of distant metastases, preparation with rhTSH stimulation may be considered as an alternative to thyroid hormone withdrawal prior to adjuvant RAI treatment.

(Weak recommendation, Low-quality evidence)

(C) In patients with ATA high-risk DTC with attendant higher risks of disease-related mortality and morbidity, more controlled data from long-term outcome studies are needed before rhTSH preparation for RAI adjuvant treatment can be recommended.

(No recommendation, Insufficient evidence)

(D) In patients with DTC of any risk level with significant comorbidity that may preclude thyroid hormone withdrawal prior to iodine RAI administration, rhTSH preparation should be considered. Significant comorbidity may include (a) a significant medical or psychiatric condition that could be acutely exacerbated by hypothyroidism, leading to a serious adverse event, or (b) inability to mount an adequate endogenous TSH response with thyroid hormone withdrawal.

(Strong recommendation, Low-quality evidence)

Recombinant human thyrotropin (trade name Thyrogen) is currently approved by many international authorities including the United States Food and Drug Administration (FDA) and Health Canada for use in preparation for RAI remnant ablation in patients who have undergone a near-total or total thyroidectomy for well-differentiated thyroid cancer and who do not have evidence of distant metastases (FDA, www.accessdata.fda.gov; Health Canada, http://webprod.hc-sc.gc.ca). Data from a compassionate use observational study suggest that rhTSH raises serum TSH measurements in patients who are unable to mount an endogenous TSH rise and appears to reduce the risk of hypothyroid-related complications in patients with significant medical or psychiatric comorbidity (697).

Some of the complications that were reported to be avoided by use of rhTSH included worsening of psychiatric illness, respiratory compromise, central nervous system (CNS) compromise, aggravation of congestive heart failure, and aggravation of coronary artery disease (697).

Multiple RCTs have focused on short-term remnant ablation outcomes in low and intermediate risk DTC with lower risk features, using rhTSH compared to thyroid hormone withdrawal. In six RCTs of patients with well-differentiated thyroid cancer without distant metastases undergoing RAI remnant ablation (T1–T3, N1 or N0, all M0), the rate of successful remnant ablation, was not significantly different after rhTSH preparation compared to thyroid hormone withdrawal, using $^{131}$I dose activities ranging from 30 to 100 mCi (691,698–702). Patients with resected cervical lymph node metastases were included in five of these trials (691,699–702), and these may be assumed to be confined to the central neck, given the extent of primary surgery described in these studies. In one trial, a further inclusion restriction was fewer than five positive nodes at the time of the primary surgery (702). Some potential limitations of the existing RCTs in this area include lack of blinding of patients and treating physicians (because it was not feasible). In five of the RCTs that examined health-related quality of life around the time of remnant ablation, this outcome was significantly worse in patients who underwent thyroid hormone withdrawal compared to rhTSH preparation, and this was attributed to hypothyroid symptoms (691,699–702). However, in three of these RCTs, which examined longer term outcomes, there is some uncertainty on the optimal level pre–RAI TSH and rate of successful remnant ablation (692–696). For example, in a secondary analysis of a RAI remnant ablation activity RCT, Fallahi et al. (692) reported that a pre-RAI TSH of >25 following (LT4 and LT3) thyroid hormone withdrawal was significantly associated with increased likelihood of successful remnant ablation (odds ratio 2.36, [95% CI 1.28–4.35], $p=0.006$), after adjustment for RAI activity, baseline serum Tg. on-LT4 TSH level, sex, age, histology, baseline RAI uptake, and extent of surgery). In two retrospective studies, each including several hundred DTC patients who underwent thyroid hormone withdrawal, no significant association was observed between pre-RAI TSH and rate of successful remnant ablation, in respective multivariate analyses adjusted for relevant variables such as disease extent, $^{131}$I activity, and sex (695,696). However, results of these two studies may not necessarily be extrapolated to TSH levels below 30 mU/L, given that patients with such TSH thresholds were not generally considered eligible for RAI ablation in these studies. Pre–RAI ablation TSH was not a significant predictor of becoming disease free without further treatment in a secondary subgroup analysis of 50 patients who underwent thyroid hormone withdrawal, but the small number of patients in this subgroup may have limited the statistical power for a multivariate analysis (694). In summary, there is some uncertainty on the optimal level pre–RAI treatment TSH following thyroid hormone withdrawal in considering long-term outcome effects.

**[B39] Can rhTSH (Thyrogen) be used as an alternative to thyroxine withdrawal for remnant ablation or adjuvant therapy in patients who have undergone near-total or total thyroidectomy?**

(A) In patients with ATA low-risk and ATA intermediate-risk DTC without extensive lymph node involvement (i.e., T1–T3, N0/Nx/N1a, M0), in whom RAI remnant ablation or adjuvant therapy is planned, preparation with rhTSH stimulation is an acceptable alternative to thyroid hormone withdrawal for achieving remnant ablation, based on evidence of superior short-term quality of life, noninferiority of remnant ablation efficacy, and multiple consistent observations suggesting no significant difference in long-term outcomes.

(Strong recommendation, Moderate-quality evidence)

(B) In patients with ATA intermediate-risk DTC who have extensive lymph node disease (multiple clinically involved LN) in the absence of distant metastases, preparation with rhTSH stimulation may be considered as an alternative to thyroid hormone withdrawal prior to adjuvant RAI treatment.

(Weak recommendation, Low-quality evidence)

(C) In patients with ATA high-risk DTC with attendant higher risks of disease-related mortality and morbidity, more controlled data from long-term outcome studies are needed before rhTSH preparation for RAI adjuvant treatment can be recommended.

(No recommendation, Insufficient evidence)

(D) In patients with DTC of any risk level with significant comorbidity that may preclude thyroid hormone withdrawal prior to iodine RAI administration, rhTSH preparation should be considered. Significant comorbidity may include (a) a significant medical or psychiatric condition that could be acutely exacerbated by hypothyroidism, leading to a serious adverse event, or (b) inability to mount an adequate endogenous TSH response with thyroid hormone withdrawal.

(Strong recommendation, Low-quality evidence)
quality of life, there was no significant difference in measurements between patients who had received rhTSH compared to those who had thyroid hormone withdrawal preparation, 3 or more months after RAI remnant ablation (699,701,702). A smaller RCT including a total of 25 individuals who had incidentally discovered PTC in the course of thyroidectomy for multinodular goiter, showed that mean Tg measurements were similar at various time points out to about 20 months in 13 individuals who were prepared for remnant ablation using rhTSH within a week after thyroidectomy, compared to those prepared by LT4 withdrawal for 4–6 weeks postoperatively (703). A meta-analysis pooling data from 1535 patients in all seven trials described herein, suggested that the rates of remnant ablation success were not significantly different using rhTSH compared to thyroid hormone withdrawal (risk ratio 0.97 [95% CI 0.94–1.01]) (704). Furthermore, a pooled analysis suggested that quality-of-life measures were superior on the day of remnant ablation in the rhTSH group, with no significant difference between groups 3 months later (704). Another meta-analysis including six of the previously mentioned RCTs, also suggested that the success of remnant ablation was not significantly different between patients prepared with rhTSH or thyroid hormone withdrawal, and this result was robust using a variety of definitions of remnant ablation success (705). In summary, the use of rhTSH for preparation for remnant ablation is associated with superior short-term quality of life and similar rates of successful remnant ablation compared to traditional thyroid hormone withdrawal.

There are some limited long-term outcome data following rhTSH preparation for RAI treatment compared to thyroid hormone withdrawal. In one aforementioned RCT in ATA low- and intermediate-risk patients (700), follow-up data were reported at a median of 3.7 years later for 51 of the original 63 patients, and rates of reoperation for cervical neck recurrence were essentially identical between groups (4% of patients), with no deaths (706). In the same study, repeat treatment with 131I for detectable Tg or imaging evidence of disease was performed in 4 of the 28 patients in the rhTSH group and 5 of the 23 patients in the hypothyroid group in this study (706). The low number of thyroid cancer–related deaths and recurrences in this small trial limit the ability to make meaningful statistical comparisons of long-term outcomes. In another RCT including 44 mixed risk–level DTC patients, who were subjected to either rhTSH preparation (n=24) for RAI treatment within a week after surgery or 4–6 weeks of LT4 withdrawal (n=20), after a mean follow-up period of about 52 months, only one individual in the rhTSH group was histologically proven to have recurrent disease (lymph nodes and bone) (707). However, this study was likely underpowered to detect meaningful differences in this outcome, and the final data analysis would be limited by lack of data available at the final follow-up (i.e., in progress) for 18% of the trial participants (707). In one prospective (694) and two retrospective (708,709) observational studies including largely ATA low- and intermediate-risk DTC patients, no significant difference was observed in the long-term presence of clinically significant disease, regardless of whether rhTSH or thyroid hormone withdrawal was used in preparing for therapeutic RAI for DTC patients. In a subgroup analysis of 183 DTC patients with level N1b nodal disease from one of the aforementioned retrospective studies, the rate of NED at last follow-up was not significantly different in patients prepared with thyroid hormone withdrawal (26.8%) compared to those prepared with rhTSH (33.7%) (708). In a smaller retrospective, multicenter study, Pitoia et al. (710) reported that in 45 consecutive Tg antibody-negative patients with T3-T4/N1-Nx/M0 disease, the absence of persistence or recurrence of disease after a mean follow-up of about 3 years was 72% in patients pretreated with rhTSH and 59% in those pretreated with thyroid hormone withdrawal (p=0.03). In a multicenter retrospective analysis of patients with T4 disease with or without nodal or distant metastases (T4, N0/N1, M0/M1), the rate of stimulated Tg <2 ng/mL Tg among antibody-negative patients was 67.7% (42 of 62) in patients prepared for RAI treatment with rhTSH, compared to 57.8% (37 of 64) in those prepared with thyroid hormone withdrawal (6-month follow-up) (711). In a retrospective analysis of 175 patients with RAI-avid metastatic disease to lungs and/or bone, the authors observed no significant difference in overall survival after a mean follow-up period of 5.5 years, among patients prepared prior to RAI treatment with rhTSH alone for all RAI treatments, thyroid hormone withdrawal for all RAI treatments, or thyroid hormone withdrawal for initial treatments followed by rhTSH for subsequent treatment(s) (712). In this study, whole-body and blood dosimetry studies were performed in all patients; therefore, the results may not be extrapolated to RAI fixed dosing. Some important differences among groups in this study that could have impacted the findings included differences in cumulative RAI activities received and longer follow-up in groups who had thyroid hormone withdrawal (712). Although the authors performed a multivariable analysis examining for predictors of overall survival, and method of TSH stimulation was not significant, this model did not adjust for these variables. In a two-center retrospective analysis comparing responses to treatment using RECIST 1.1 criteria in 56 patients with distant metastatic disease pretreated with either rhTSH or thyroid hormone withdrawal prior to RAI, there were no differences in outcomes between groups after a mean follow-up period of about 6 years (713). Also in this study, there were important baseline differences among groups, such as rates of use of dosimetry and mean cumulative RAI activity. Rates of xerostomia, leukopenia, or thrombocytopenia were not significantly different between treatment groups in this study. The overall mortality rate was 20% in the rhTSH group (3 of 15) and 7.3% in the thyroid hormone withdrawal group (3 of 41) (p=0.188) (713), although the study was likely not sufficiently large to examine differences in this important outcome. The findings of the latter study cannot be readily extrapolated to fixed-dosing RAI treatment regimens because 80% of the individuals in the rhTSH group and 46% in the thyroid hormone withdrawal group had dosimetry-based RAI treatment. RCTs comparing rhTSH to thyroid hormone withdrawal preparation pre–RAI treatment, are clearly needed to guide clinical care in higher risk DTC patients.

[B40] What activity of 131I should be used for remnant ablation or adjuvant therapy?

- **RECOMMENDATION 55**

(A) If RAI remnant ablation is performed after total thyroidectomy for ATA low-risk thyroid cancer or intermediate-risk disease with lower risk features (i.e., low-
volume central neck nodal metastases with no other known gross residual disease or any other adverse features), a low administered activity of approximately of 30 mCi is generally favored over higher administered activities.

(Strong recommendation, High-quality evidence)

(B) Higher administered activities may need to be considered for patients receiving less than a total or near-total thyroidectomy in which a larger remnant is suspected or in which adjuvant therapy is intended.

(Weak recommendation, Low-quality evidence)

**RECOMMENDATION 56**

When RAI is intended for initial adjuvant therapy to treat suspected microscopic residual disease, administered activities above those used for remnant ablation up to 150 mCi are generally recommended (in absence of known distant metastases). It is uncertain whether routine use of higher administered activities (>150 mCi) in this setting will reduce structural disease recurrence for T3 and N1 disease.

(Weak recommendation, Low-quality evidence)

Successful remnant ablation can be defined by an undetectable stimulated serum Tg, in the absence of interfering Tg antibodies, with or without confirmatory nuclear or other imaging studies. An alternative definition in cases in which Tg antibodies are present is the absence of visible RAI uptake on a subsequent diagnostic RAI scan. In this section, we only included published original RCTs or systematic reviews/meta-analyses of such studies examining the impact of various $^{131}$I dose activities on the rate of successful remnant ablation or thyroid cancer outcomes (including thyroid cancer–related deaths or recurrences) in adult patients with well-differentiated thyroid carcinoma who had been treated with total or near-total thyroidectomy and who were not known to have any gross residual disease following surgery. Our search yielded six RCTs (692,699,701,714–716), the majority of which had had no blinding of patients and health care providers (699,701,714–717).

In the trial by Fallahi et al. (692), the randomization program was prepared and executed by a technologist in the laboratory where the $^{131}$I activities were prepared for dis-amping in coded vials, without revealing the administered activity to participants and health care providers. The pathologic stage of disease of patients included was TNM stage pT1 to pT3 in four trials (699,714,715,717), whereas only pT1 or pT2 patients were eligible in one trial (701), and primary tumor size or tumor size staging was not reported in two trials (692,716). Some patients with known small-volume lymph node disease were included in five of the trials (699,701,714,715,717), but patients with known lymph node disease were excluded from one trial (692), and lymph node staging was not reported in another trial (716). Although the specific levels and size of lymph node metastases at baseline were not clearly reported, data reported on surgical extent suggested that these were consistent with relatively nonbulky central neck nodal metastases resected in the course of thyroidectomy with or without central neck dissection (699,701, 714,715,717).

The $^{131}$I activities compared were as follows: 30 mCi compared to 100 mCi in four trials (692,699,701,714), 50 mCi compared to 100 mCi in two trials (715,716), or 30 mCi compared to 60 mCi or 100 mCi in one study comprising pooled long-term outcome data from two staged smaller trials from the same group (717).

The rate of successful remnant ablation was reported to be not inferior using an administered activity of 30 mCi as compared to 100 mCi in three trials after preparation with thyroid hormone withdrawal (699,701,714) and in two trials after preparation with rhTSH (699,701), including data from the two large factorial design trials in both comparisons (699,701). Pilli et al. (715) reported that an administered activity of 50 mCi was not inferior to 100 mCi in achieving successful remnant ablation, following preparation with rhTSH. In contrast, Zaman et al. (716) suggested that 100 mCi may be superior to 50 mCi following thyroid hormone withdrawal, but the small size (40 patients) and lack of reporting of statistical comparisons are important limitations of this study. The third largest trial (341 patients randomized) by Fallahi et al. (692) reported that an administered activity of 100 mCi was superior to 30 mCi in achieving successful remnant ablation after thyroid hormone withdrawal. The rate of initial successful remnant ablation (as defined by the primary authors) was highly variable among trials, and the following rates were reported following initial administration of 100 mCi of $^{131}$I: 64% in the trial from Fallahi et al. (692), 56% in the trial from Maenpaa et al. (714), 89% in the trial from Mallick et al. (699), 67% in the trial from Pilli et al. (715), 94% in the trial from Schlumberger et al. (701), and 60% in the trial from Zaman et al. (716). The reasons explaining variability in the rates of successful RAI remnant ablation among trials are not completely understood but could potentially be due to differences in study populations, completeness of surgery (including size of the remaining remnant), or sensitivity of techniques used to evaluate outcomes (such as Tg assays or diagnostic imaging studies).

Short-term side effects in the weeks following remnant ablation have been reported to be more frequent in patients treated with 100 mCi as compared to 30 mCi activities by Mallick et al. (699), and a similar trend was reported by Maenpaa et al. (714). Repeat treatment with additional $^{131}$I has been reported to be more frequent in patients treated with 30 mCi as compared to higher activities in three trials (692,699,717), but not in one trial (714). Long-term outcome data from randomized trials in this area are limited by relatively low event rates, potentially underpowering statistical analyses. Kukulski et al. (717) followed 390 patients that had been randomized to either 30, 60, or 100 mCi administered activities for remnant ablation, and reported the following event rates after a median period of 10 years following treatment: local relapse in 2% following an initial 30 mCi activity compared to 3% for initial administered activities of 60 or 100 mCi (reported to be not significantly different), and distant metastatic recurrence in 0% of the patients in all of the treatment groups. Maenpaa et al. (714) followed 160 patients who were randomized to either 30 or 100 mCi for a median of 51 months, and they reported the following outcomes: reoperation for resection of thyroid cancer in lymph nodes in 7% and 8% of patients who were treated with activities of 30 and 100 mCi, respectively; distant metastatic recurrence in 0% and 4% of patients treated with activities of 30 and 100 mCi, respectively; and no thyroid cancer–related deaths in any of the treatment groups.
Overall, the rate of successful remnant ablation in patients who have undergone total or near-total thyroidectomy appears to be not inferior in patients treated with 30 mCi compared to 100 mCi in the majority of studies comparing these activities and particularly in studies achieving the highest successful ablation rates. Rates of short-term adverse effects may be higher after administration of 100 mCi $^{131}$I compared to 30 mCi, in a small number of trials examining these outcomes. Limited long-term RCT data on the impact of various activities for remnant ablation or adjuvant therapy are available, but thyroid cancer–related recurrences or deaths do not appear to be higher with the use of lower initial activities for remnant ablation, compared to higher administered activities. Four recent systematic reviews and meta-analyses reported results that are supportive of these conclusions (718–721), although some of the predefined study inclusion criteria in these reviews were generally not as strict as defined in our review (particularly for variables such as the extent of primary surgery or the stringency of Tg threshold in the definition of success of remnant ablation); it is also important to note that in some of these meta-analyses, statistically significant heterogeneity (variability) of treatment effect was noted for the pooled analyses on successful remnant ablation (718,720,721). A recent retrospective database analysis by Verburg et al. (722) with longer follow-up than most earlier studies adds a note of caution for the use of lower administered activities in older low-risk patients. They followed 698 patients with low-risk DTC (pT1 or pT2 and no involved LN) for at least 5 years. There were no long-term (10–15 year) overall survival or disease-specific survival differences in younger patients (<45 years old) who received lower administered activities of $^{131}$I ($\leq$54 mCi) compared with those receiving higher administered activities. The older patients (≥45 years old), however, who received lower administered activities of $^{131}$I ($\leq$54 mCi) did have a lower disease-specific survival compared with those receiving higher administered activities. Disease was defined as abnormal structural or functional imaging or a detectable serum Tg after TH stimulation. The absolute disease-specific survival remained high even in the patients receiving lower administered activities of $^{131}$I, and there were no differences in overall survival in these older patients.

In 2009, the ATA guidelines task force recommended a fixed administered activity of between 100 and 200 mCi for adjuvant RAI treatment if residual microscopic disease is suspected or if an aggressive histologic variant of DTC was present (25). Since that time, at least five retrospective, single institution studies have compared clinical outcomes following various adjuvant RAI fixed activities in ATA intermediate-risk and ATA higher risk patients, without distant metastases (723–726). In comparing rates of disease persistence or recurrence in 225 ATA intermediate-risk DTC patients treated with 30 to 50 mCi compared to ≥100 mCi of adjuvant RAI, Castagna et al. (723) reported no significant difference in rates of successful remnant ablation or in long-term disease persistence/recurrence between groups. However, there were some statistically significant differences between the treatment groups that may have influenced these results, including higher numbers of men and individuals with lateral neck nodal disease and a longer follow-up period (which may increase the rate of detection) in the higher administered activity group of this study (723). In another study including 176 DTC patients with a primary tumor size ≤2 cm in diameter and microscopic extrathyroidal extension, no significant differences were found in a comparison of rates of successful remnant ablation and long-term recurrences in patients treated with 30 mCi $^{131}$I compared with 149 mCi (724). In this study, no recurrences were noted in either group after a median follow-up of 7.2 years. Although the mean primary tumor size was higher in the group treated with higher activities compared with the lower activity group in this study ($p<0.001$), the difference in mean tumor diameter was only 2 mm, so it may be of questionable clinical significance (724). Kruijff et al. (725) reported the results of multiple secondary subgroup analyses on data from 341 patients with T3 PTC, in which a postsurgical $^{131}$I administered activity of ≤75 mCi was compared to >75 mCi. In this study, the respective rates of disease recurrence, mortality, and stimulated Tg ≥2 ng/mL were not significantly different in the lower administered activity group (i.e., 7%, 3%, 72%), compared to the higher activity group (12%, 1.7%, 64%) (respective $p$-values of 0.55, 0.43, 0.40). Furthermore, in this study, a multivariate analysis using data from 1171 mixed-risk DTC patients without distant metastases suggested that there was no significant difference in risk of disease recurrence with the use of >75 mCi of $^{131}$I postoperatively compared to ≤75 mCi (higher administered activity hazard ratio 1.57 ([95% CI 0.61–3.98], $p=0.341$), after adjustment for age, sex, primary tumor size, presence of vascular invasion, multifocality, and lymph node positivity, with a mean follow-up period of 60 months) (725). In another study comparing rates of disease structural recurrence/persistence in 181 patients with positive N1b lymph nodes, Sabra et al. (726) reported no statistically significant difference between the following approximate fixed administered activity categories: 75–139 mCi with a median of 102 mCi (31%), 140–169 mCi with a median of 150 mCi (32%), and 170–468 mCi with a median of 202 mCi (23%) ($p=0.17$). Consistent with this finding, no significant correlation between RAI activity and best clinical response was observed in this study. In respective secondary subgroup analyses, a dose response was apparent in individuals ≥45 years of age, but not younger individuals, and the authors cautioned about the use of fixed RAI activities exceeding 150 mCi because of concerns about potential toxicity in the context of renal impairment (726). In the three studies utilizing either thyroid hormone withdrawal or rhTSH for RAI treatment preparation (723,725,726), insufficient data were reported to make any meaningful interpretation on any relationship between administered activity in the context of preparation method. In one study, RAI adjuvant treatment was performed postoperatively, presumably without initiation of thyroid hormone because rhTSH was not reported to be used (724). None of the aforementioned studies (723–726) reported on RAI toxicity or quality of life outcomes; furthermore, T4 disease was not included in these studies. Overall, there is little evidence to suggest that increasing administered activities of adjuvant RAI is necessarily associated with improvement of clinical outcomes for patients with ATA intermediate- and high-risk disease without evidence of persistent disease. There is an important unmet need for RCTs examining thyroid cancer–related outcomes, quality of life, and toxicities in patients with ATA intermediate or higher level thyroid cancer, in the absence of gross residual disease or distant metastases.
[B41] Is a low-iodine diet necessary before remnant ablation?

**RECOMMENDATION 57**

A low iodine diet (LID) for approximately 1–2 weeks should be considered for patients undergoing RAI remnant ablation or treatment.

(Weak recommendation, Low-quality evidence)

It is important for health care providers to inquire about a history of possible high-dose iodine exposure (e.g., IV contrast, amiodarone, or others) in considering timing of scheduling RAI therapy or imaging. There are no studies examining whether the use of a LID in preparation for RAI remnant ablation or treatment impacts long-term disease-related recurrence or mortality rates. In a recent systematic review of observational studies in this area, the most commonly studied LIDs allowed for ≤50 μg/d of iodine for 1–2 weeks and that the use of LIDs appeared to be associated with reduction in urinary iodine excretion as well as increase in $^{131}$I uptake, compared to no LID (727). There is conflicting evidence on the impact of a LID on the outcome of remnant ablation success (727), with the best available evidence largely restricted to retrospective analyses, using historical controls (728,729). In a study including a total of 120 patients, the use of a 4-day LID (with seafood restriction for 1 week) was associated with a higher rate of remnant ablation success (defined by absent neck activity and stimulated Tg week) was associated with a higher rate of remnant ablation (727), with the best available evidence largely restricted to retrospective analyses, using historical controls (728,729). In a study including a total of 94 patients, comparing a more stringent LID to a less stringent diet of restricted salt/vitamins/seafood, each for 10 to 14 days, there was no significant difference in rate of successful remnant ablation, using a visually negative $^{131}$I scan to define that outcome (729). The optimal stringency and duration of LID (if any) prior to therapeutic RAI is not known. In a RCT including 46 patients, the increase in $^{131}$I uptake and reduction in urinary iodine excretion was not significantly different between patients who followed a LID for 2 weeks compared to 3 weeks prior to RAI scanning (730), suggesting that there may be little reason to extend the LID past 2 weeks. However, a lack of association between urinary iodine excretion and rate of successful thyroid ablation has been reported in patients not specifically prescribed a LID (731); the absence of a specific LID comparison group in this study may limit the generalizability of the findings to situations in which a specific LID is prescribed. Such findings may suggest, however, that the routine measurement of urinary iodine excretion, outside of possibly a research setting or suspected iodine contamination, may not be necessary. Although LIDs may be cumbersome or unpalatable, serious side effects are relatively infrequent (727), with case reports of potentially life-threatening hyponatremia occurring largely in elderly patients who were subject to thyroid hormone withdrawal, often in the presence of metastatic disease, sometimes concurrently treated with thiazide diuretics, and with a LID duration of longer than a week in the majority of cases (732). It is important to avoid restriction of noniodized salt during the LID, since this may be associated with hyponatremia, especially in patients undergoing thyroid hormone withdrawal. Some examples of LID descriptions for patients, may be found at the following websites: ATA (www.thyroid.org/faq-low-iodine-diet/), ThyCa: Thyroid Cancer Survivors’ Association, Inc. (http://thyca.org/rai.htm#diet), Light of Life Foundation (www.checkyourneck.com/About-Thyroid-Cancer/Low-Iodine-Cookbook), and Thyroid Cancer Canada (www.thyroidcancernet.org/userfiles/files/LID_English_Aug_2013_final.pdf).

[RECOMMENDATION 58]

A posttherapy WBS (with or without SPECT/CT) is recommended after RAI remnant ablation or treatment, to inform disease staging and document the RAI avidity of any structural disease.

(Strong recommendation, Low-quality evidence)

The literature on the utility of posttherapy RAI scans is largely based on single-center retrospective studies (643,733–736), many of which included a relatively large proportion of ATA intermediate- and high-risk DTC patients (643,733,734). In a comparison of the results of pretherapy $^{131}$I scans to posttherapy scans, the rate of newly discovered lesions on posttherapy scans was reported to be between 6% and 13%. In a study examining post–remnant ablation scans in 60 DTC patients, the disease stage was altered in 8.3% of individuals (735). In older literature, it had been reported that posttherapy scanning demonstrated new findings in 31% of 39 cases studied, but the detection of thyroid foci was included in that outcome, whereas almost a third of the patients (12 of 39) had a sizeable portion of their thyroid remaining following primary surgery (736). In the recent posttherapy scan literature, the $^{131}$I activities ranged from 30 to 300 mCi (733–735), and the timing of scans was between 2 and 12 days following therapeutic RAI (643,733–735,737,738), with some studies prescribing a preparatory LID (643,735,738) and others not prescribing a LID (733,734). In one study, posttherapy scan images were compared on the third and seventh day following ablative or therapeutic RAI administration for mixed-risk DTC (following thyroid hormone withdrawal) (738). The authors of this study reported that the concordance of lesions detected on both scans was 80.5% (108 of 135 patients), with 7.5% of early scans providing more information than late scans, and 12% of late scans providing more information than early scans (738). A limitation in interpreting the posttherapy scan literature is that all of the lesions identified on posttherapy scans were not always confirmed to represent structural disease (i.e., using cross-sectional imaging, histopathology, or long-term outcome data), and readers of posttherapy scans were generally not specifically blinded to the results of other investigations, such as pretherapy RAI scans.

The potential utility of the combination of RAI posttherapy scanning in conjunction with SPECT/CT has been examined in multiple prospective (737,739,740) and retrospective studies (741–743). The majority of these studies (737,739–742) have independently confirmed the presence of disease by means such as alternative imaging studies, histopathology, or clinical follow-up. In a single-center...
In another study, the routine addition of neck/chest nodes to the treatment plan of patients with "tiny" lung nodules in 19 patients, mediastinal lymph nodes, and osteolytic bone metastases in three patients (743). In situations in which SPECT/CT may not be feasible to perform in conjunction with posttherapy RAI scanning, the SPECT/CT portion identified non–iodineavid lesions in 22% of patients (32 of 148), although the underlying pathologic diagnosis or long-term clinical significance of these lesions was not clearly reported (i.e., underlying pathologic diagnosis or long-term clinical significance of these lesions was not clearly reported). In one study examining the use of routine cervical/thoracic SPECT/CT in conjunction with posttherapy RAI scanning, the SPECT/CT portion identified non–iodineavid lesions in 22% of patients (32 of 148), although the underlying pathologic diagnosis or long-term clinical significance of these lesions was not clearly reported (i.e., "tiny" lung nodules in 19 patients, mediastinal lymph nodes <5 mm in 10 patients, and osteolytic bone metastases in three patients) (743). In situations in which SPECT/CT may not be feasible to perform in conjunction with posttherapy RAI scan, clinical judgment needs to prevail on the utility of alternative cross-sectional imaging studies, considering factors such as clinical-pathologic stage, the completeness of surgery, inappropriate thyroglobulinemia, and, if performed, the posttherapy RAI scan result.

[B43] Early management of DTC after initial therapy

[B44] What is the appropriate degree of initial TSH suppression?

**RECOMMENDATION 59**

(A) For high-risk thyroid cancer patients, initial TSH suppression to below 0.1 mU/L is recommended.

(Strong recommendation, Moderate-quality evidence)

(B) For intermediate-risk thyroid cancer patients, initial TSH suppression to 0.1–0.5 mU/L is recommended.

(Weak recommendation, Low-quality evidence)

(C) For low-risk patients who have undergone remnant ablation and have undetectable serum Tg levels, TSH may be maintained at the lower end of the reference range (0.5–2 mU/L) while continuing surveillance for recurrence. Similar recommendations hold for low-risk patients who have not undergone remnant ablation and have undetectable serum Tg levels.

(Weak recommendation, Low-quality evidence)

(D) For low-risk patients who have undergone remnant ablation and have low-level serum Tg levels, TSH may be maintained at or slightly below the lower limit of normal (0.1–0.5 mU/L) while surveillance for recurrence is continued. Similar recommendations hold for low-risk patients who have not undergone remnant ablation, although serum Tg levels may be measurably higher and continued surveillance for recurrence applies.

(Weak recommendation, Low-quality evidence)

(E) For low-risk patients who have undergone lobectomy, TSH may be maintained in the mid to lower reference range (0.5–2 mU/L) while surveillance for recurrence is continued. Thyroid hormone therapy may not be needed if patients can maintain their serum TSH in this target range.

(Weak recommendation, Low-quality evidence)

DTC expresses the TSH receptor on the cell membrane and responds to TSH stimulation by increasing the expression of several thyroid specific proteins (Tg, sodium-iodide symporter) and by increasing the rates of cell growth (744). Suppression of TSH, using supraphysiologic doses of LT₄, is used commonly to treat patients with thyroid cancer in an effort to decrease the risk of recurrence (275,671,745–747). A meta-analysis supported the efficacy of TSH suppression therapy in preventing major adverse clinical events (RR = 0.73 [CI = 0.60–0.88], p < 0.05) (745). A large RCT from Japan (748) showed that disease-free survival was equivalent in patients with normal TSH (0.4–5 mU/L) compared with those on LT₄ suppression therapy (TSH <0.01 mU/L). Extent of residual disease is uncertain in these patients in that most did not undergo total thyroidectomy or RAI ablation and Tg levels were not monitored or reported, making direct comparisons to a North American approach difficult.

Retrospective and prospective studies have demonstrated that TSH suppression to below 0.1 mU/L may improve outcomes in high-risk thyroid cancer patients (275,749), though no such evidence of benefit has been documented in low-risk patients. Higher degrees of suppression to <0.03 mU/L may offer no additional benefit (746). A prospective, nonrandomized cohort study (671) of 2936 patients found that overall survival improved significantly when the TSH was suppressed to undetectable levels in patients with NTCTCSG stage III or IV disease and suppressed to the subnormal to undetectable range in patients with NTCTCSG stage II disease; however, in the latter group there was no incremental benefit from suppressing TSH to undetectable levels. Patients in the NTCTCSG stage II classification are somewhat different from AJCC/UICC stage II patients. Suppression of TSH was not beneficial in patients with NTCTCSG stage I disease. In another study, there was a positive association between serum TSH levels and the risk for recurrent disease (impacted management in 35% [8 of 23] of such patients in another study) (741). The incremental value of SPECT/CT in impacting therapeutic planning appears to be greatest in studies in which its use was reserved for situations of posttherapy scan diagnostic uncertainty (impacted therapy in 24.4% [8 of 33] of cases, all of which went on to surgery) (739), or when disease was advanced and WBS was inconclusive (impacted management in 35% [8 of 23] of such patients in another study) (740). The routine addition of neck/chest SPECT/CT to all posttherapy scans was estimated to alter postsurgical ATA recurrence risk estimate in 6.4% (7 of 109) of patients (743), impact therapeutic planning in about 2% of cases (742), and reduce the need for additional cross-sectional imaging in 20% of cases (29 of 148). In one study examining the use of routine cervical/thoracic SPECT/CT in conjunction with posttherapy RAI scanning, the SPECT/CT portion identified non–iodineavid lesions in 22% of patients (32 of 148), although the underlying pathologic diagnosis or long-term clinical significance of these lesions was not clearly reported (i.e., “tiny” lung nodules in 19 patients, mediastinal lymph nodes <5 mm in 10 patients, and osteolytic bone metastases in three patients) (743). In situations in which SPECT/CT may not be feasible to perform in conjunction with posttherapy RAI scan, clinical judgment needs to prevail on the utility of alternative cross-sectional imaging studies, considering factors such as clinical-pathologic stage, the completeness of surgery, inappropriate thyroglobulinemia, and, if performed, the posttherapy RAI scan result.

TSH suppression to below 0.1 mU/L is recommended. (Weak recommendation, Low-quality evidence)
disease and cancer-related mortality (750). Adverse effects of TSH suppression may include the known consequences of subclinical thyrotoxicosis, including exacerbation of angina in patients with ischemic heart disease, increased risk for atrial fibrillation in older patients (751), and increased risk of osteoporosis in postmenopausal women (748,752–754). Therefore, optimal TSH goals for individual patients must balance the potential benefit of TSH suppression with the possible harm from subclinical thyrotoxicosis especially in patients with medical conditions that can be exacerbated with aggressive TSH suppression.

There is little evidence to guide TSH targets or the use of thyroid hormone in ATA low-risk patients who have undergone lobectomy. Most of the studies evaluating lobectomy for these patients do not discuss TSH targets or the use of thyroid hormone therapy or note that these data were unavailable in the databases studied (318,323–327). Vaisman et al. (328) noted that “levothyroxine was often not given after lobectomy if the patient maintained thyroid function tests within the reference range,” while Matsuzaka et al. (322) noted that “TSH suppression therapy was performed in most of the cases postoperatively, but the patients’ TSH levels were not analyzed in this study.” A recent study by Ebina et al. (755) retrospectively analyzed low-risk patients who had undergone lobectomy and had not received thyroid hormone therapy. After a mean follow-up of 8.3 years, only 13% of the 674 patients undergoing lobectomy became overtly hypothyroid. The 10-year cause-specific and disease-specific survivals were not different between the patients who had undergone thyroid function tests within the reference range, and those who had not. A closer look at patients who underwent thyroidectomy versus a lesser operation, though an increased incidence of second tumors in thyroid origin (760) and could be considered for patients with locally advanced disease undergoing external beam radiation therapy. It is unproven whether patients with rising Tg in the setting of no identifiable progression of anatomical disease have sufficiently high future risks from disease to justify the application of adjuvant systemic therapy beyond RAI and/or TSH suppressive therapy using LT₄.

[C1] DTC: LONG-TERM MANAGEMENT AND ADVANCED CANCER MANAGEMENT GUIDELINES

[B47] Systemic adjuvant therapy

■ RECOMMENDATION 61

There is no role for routine systemic adjuvant therapy in patients with DTC (beyond RAI and/or TSH suppressive therapy using LT₄).

(Strong recommendation, Low-quality evidence)

There are no clinical trial data to indicate that any adjuvant therapy beyond RAI and/or TSH suppressive therapy using LT₄ has a net beneficial role in DTC patients. Furthermore, as the prognosis of DTC patients in complete remission and without any indication of active systemic disease is very good—and as toxicities, and even the risk of death, from use of kinase inhibitor therapies are appreciable—toxicities/risks have strong potential to exceed expected therapeutic benefit in the adjuvant context in most patients with DTC.

Whether populations of DTC patients might be identifiable who have sufficiently great future risks from recurrent disease to justify the corresponding risks attendant to the application of adjuvant systemic therapy (beyond RAI and/or TSH suppressive therapy using LT₄) remains uncertain. Doxorubicin may act as a radiation sensitizer in some tumors of thyroid origin (760) and could be considered for patients with locally advanced disease undergoing external beam radiation therapy. It is unproven whether patients with rising Tg in the setting of no identifiable progression of anatomical disease have sufficiently high future risks from disease to justify the application of adjuvant systemic therapy beyond RAI and/or TSH suppressive therapy using LT₄.

[C2] What are the appropriate features of long-term management?

Accurate surveillance for possible recurrence in patients thought to be free of disease is a major goal of long-term follow-up. Tests with high specificity allow identification of patients unlikely to experience disease recurrence, so that less aggressive management strategies can be used that may be more cost effective and safe. Similarly, patients with a higher risk of recurrence are monitored more aggressively because it is believed that early detection of recurrent disease offers the best opportunity for effective treatment. A large study (761) found that the residual life span in disease-free patients treated with total or near-total thyroidectomy, ¹³¹I for remnant ablation, and in some cases high-dose ¹³¹I for residual disease was similar to that in the general Dutch population. In contrast, the life expectancy for patients with persistent disease was reduced to 60% of that in the general population but varied widely depending upon tumor features. Age was not a factor in disease-specific mortality in a comparison of patients with age-matched individuals in the Dutch population. Treatment thus appears safe and does not shorten life expectancy. Although an increased incidence of second tumors in thyroid
cancer patients has been recognized after the administration of high cumulative activities of $^{131}$I (762,763), this elevated risk was not found to be associated with the use of $^{131}$I in another study (764). RAI therapy in low-risk patients did not affect median overall survival in one study (765), but did increase second primary malignancies in another study (766). This risk of second primary malignancies after RAI therapy is discussed in more detail in section C33. Patients with persistent or recurrent disease are offered treatment to cure or to delay future morbidity or mortality. In the absence of such options, therapies to palliate by substantially reducing tumor burden or preventing tumor growth are utilized, with special attention paid to tumors threatening critical structures.

A second goal of long-term follow-up is to monitor thyroxine suppression or replacement therapy to avoid under-replacement or overly aggressive therapy (767).

**[C3] What are the criteria for absence of persistent tumor (excellent response)?**

In patients who have undergone total or near-total thyroidectomy and RAI treatment (remnant ablation, adjuvant therapy or therapy), disease-free status comprises all of the following (summarized in Table 13):

1. No clinical evidence of tumor
2. No imaging evidence of tumor by RAI imaging (no uptake outside the thyroid bed on the initial posttreatment WBS if performed, or if uptake outside the thyroid bed had been present, no imaging evidence of tumor on a recent diagnostic or posttherapy WBS) and/or neck US
3. Low serum Tg levels during TSH suppression (Tg <0.2 ng/mL) or after stimulation (Tg <1 ng/mL) in the absence of interfering antibodies

**[C4] What are the appropriate methods for following patients after initial therapy?**

**[C5] What is the role of serum Tg measurement in the follow-up of DTC?**

**RECOMMENDATION 62**

(A) Serum Tg should be measured by an assay that is calibrated against the CRM457 standard. Thyroglobulin antibodies should be quantitatively assessed with every measurement of serum Tg. Ideally, serum Tg and anti-Tg antibodies should be assessed longitudinally in the same laboratory and using the same assay for a given patient.

*(Strong recommendation, High-quality evidence)*

(B) During initial follow-up, serum Tg on thyroxine therapy should be measured every 6–12 months. More frequent Tg measurements may be appropriate for ATA high-risk patients.

*(Strong recommendation, Moderate-quality evidence)*

(C) In ATA low- and intermediate-risk patients that achieve an excellent response to therapy, the utility of subsequent Tg testing is not established. The time interval between serum Tg measurements can be lengthened to at least 12–24 months.

*(Weak recommendation, Low-quality evidence)*

(D) Serum TSH should be measured at least every 12 months in all patients on thyroid hormone therapy.

*(Strong recommendation, Low-quality evidence)*

(E) ATA high-risk patients (regardless of response to therapy) and all patients with biochemical incomplete, structural incomplete, or indeterminate response should continue to have Tg measured at least every 6–12 months for several years.

*(Weak recommendation, Low-quality evidence)*

**RECOMMENDATION 63**

(A) In ATA low-risk and intermediate-risk patients who have had remnant ablation or adjuvant therapy and negative cervical US, serum Tg should be measured at 6–18 months on thyroxine therapy with a sensitive Tg assay (<0.2 ng/mL) or after TSH stimulation to verify absence of disease (excellent response).

*(Strong recommendation, Moderate-quality evidence)*

(B) Repeat TSH-stimulated Tg testing is not recommended for low- and intermediate-risk patients with an excellent response to therapy.

*(Weak recommendation, Low-quality evidence)*

(C) Subsequent TSH-stimulated Tg testing may be considered in patients with an indeterminate, biochemical incomplete, or structural incomplete response following either additional therapies or a spontaneous decline in Tg values on thyroid hormone therapy over time in order to reassess response to therapy.

*(Weak recommendation, Low-quality evidence)*

Subsequent stimulated testing is rarely needed for those with NED, because there are rarely benefits seen in this patient population from repeated TSH-stimulated Tg testing (590,594,597,768). The use of sensitive methods for serum Tg may obviate the need for rhTSH stimulation in low-risk patients with a Tg on LT$_4$ treatment below 0.1–0.2 ng/mL (393,587,595,601,606,769).

**[C6] Serum Tg measurement and clinical utility**

Measurement of serum Tg levels is an important modality to monitor patients for residual or recurrent disease. Most laboratories currently use immunometric assays to measure serum Tg, and it is important that these assays are calibrated against the CRM-457 international standard. Despite improvements in standardization of Tg assays, there is still a two-fold difference between some assays (316,770), leading to the recommendation that measurements in individual patients be performed with the same assay over time. Immunometric assays are prone to interference from anti-Tg autoantibodies, which commonly cause falsely low serum Tg measurements. Moreover, variability in Tg autoantibody assays may result in falsely negative antibody levels associated with a misleadingly undetectable serum Tg due to the antibodies that are present but not detected (771). Assays for anti-Tg autoantibodies suffer from a similar variance and lack of concordance as do Tg assays (608,772), and both Tg and Tg autoantibody assays may be affected by heterophilic antibodies (773,774). The presence of Tg autoantibodies should be suspected when the surgical pathology indicates the presence of background Hashimoto thyroiditis (775). While there
is no method that reliably eliminates Tg antibody interference, radioimmunooassays for Tg may be less prone to antibody interference, which can occasionally result in falsely elevated Tg levels (776–778). However, radioimmunooassays for Tg are not as widely available, may be less sensitive than immunometric assays in detecting small amounts of residual tumor, and their role in the clinical care of patients is uncertain. In the absence of antibody interference, serum Tg has a high degree of sensitivity and specificity to detect thyroid cancer, especially after total thyroidectomy and remnant ablation. In patients with low risk for recurrence, serum Tg measurement at the time of remnant ablation/adjuvant therapy may be useful for prediction of subsequent disease-free status (605).

Most data come from studies using methods with a functional sensitivity of 1 ng/mL. Functional sensitivity of many contemporary assays is ≤0.1 ng/mL, which may lead to greater reliance of Tg on thyroid hormone therapy instead of performing Tg determination following TSH stimulation. Because TSH stimulation generally increases basal serum Tg by 5- to 10-fold, significant serum Tg levels (>1–2 ng/mL) found after TSH stimulation when using an assay with a functional sensitivity of 0.5–1 ng/mL may already be predicted in patients on LT4 treatment without TSH stimulation by a highly sensitive Tg assay when Tg levels are above 0.2 ng/mL.

The highest degrees of sensitivity for serum Tg are noted following thyroid hormone withdrawal or stimulation using rhTSH (779). Serum Tg measurements obtained during thyroid hormone suppression of TSH and less commonly following TSH stimulation may fail to identify patients with relatively small amounts of residual tumor (583,649, 780,781). These minimal amounts of residual disease are often located in the neck, and performing neck US in these patients offers the best opportunity to recognize or exclude neoplastic disease even when serum Tg is undetectable (297,782,783). Conversely, even TSH-stimulated Tg measurement may fail to identify patients with clinically significant tumor because of anti-Tg antibodies or less commonly because of defective or absent production and secretion of immunoreactive Tg by tumor cells (649,780). Tg levels should be interpreted in light of the pretest probability of clinically significant residual tumor. An aggressive or poorly differentiated tumor may be present despite low basal or stimulated Tg; in contrast, a minimally elevated stimulated Tg may occur in patients at low risk for clinically significant morbidity (784). Nevertheless, a single rhTSH-stimulated serum Tg <0.5–1.0 ng/mL in the absence of anti-Tg antibody has an approximately 98%–99.5% likelihood of identifying patients completely free of tumor on follow-up (590,591,593,597,768). Repeating rhTSH-stimulated Tg measurements may not be necessary in most cases when surveillance includes an undetectable basal serum Tg and negative ultrasonography (604,617). However, 0.5%–3% of patients may manifest clinical or biochemical recurrence in spite of an initial rhTSH-stimulated Tg of <0.5 ng/mL (592).

Initial follow-up for low-risk patients (about 85% of postoperative patients) who have undergone total or near-total thyroidectomy and 131I remnant ablation should be based mainly on TSH-suppressed Tg and cervical US, followed by TSH-stimulated serum Tg measurements if the TSH-suppressed Tg testing is undetectable (583,785). However, a Tg assay with a functional sensitivity of 0.1–0.2 ng/mL may reduce the need to perform TSH-stimulated Tg measurements during the initial and long-term follow-up of some patients. In one study using such an assay, a Tg suppressed serum Tg <0.1 ng/mL was only rarely (2%) associated with an rhTSH-stimulated Tg >2 ng/mL; however, 42% of the patients had baseline rhTSH-stimulated Tg elevation >0.1 ng/mL, but only one patient was found to have residual tumor (606). In another study using the same assay (787), a TSH-suppressed serum Tg level was >0.1 ng/mL in 14% of patients, but the false-positive rate was 35% using an rhTSH-stimulated Tg cutoff of >2 ng/mL, raising the possibility of unnecessary testing and treatment. In low-risk patients not undergoing ablation, an ultrasensitive Tg was <1 ng/mL in 91% and <2 ng/mL in 96% of patients at 9 months after thyroidectomy (644). In a second-generation assay, a cutoff of 0.15 ng/mL was shown to have a NPV of 98.6% and 91% specificity for residual disease or potential recurrence (587). The only prospective study also documented increased sensitivity of detection of disease at the expense of reduced specificity (770), and receiver operating curves have shown that a Tg level on thyroid hormone therapy around 0.2–0.3 ng/mL portends the best sensitivity and specificity for detecting persistent disease. With the use of these sensitive Tg assays, it was concluded that an annual serum Tg on LT4 treatment with periodic neck US is adequate for detection of recurrence without need for rhTSH stimulation testing in those patients with a serum Tg <0.2–0.3 ng/mL on thyroid hormone therapy (606). In patients at low to intermediate risk of recurrence, the utility of an undetectable postoperative nonstimulated Tg level is uncertain and may depend upon the functional sensitivity of the Tg assay, with some studies (632,646) observing RAI-avid metastatic foci (usually in neck lymph nodes) in 8.5%–12% of such patients, while another study (630) noted negative scans in 63 of 63 patients when the baseline Tg was <0.2 ng/mL. The different results likely relate to both the degree of intermediate versus higher risk patients in the respective cohorts, the amount of residual thyroid tissue, the elapsed time since surgery, the cutoff for functional sensitivity of the Tg assays, as well as the sensitivity of the post-RAI imaging techniques.

Approximately 20% of patients who are clinically free of disease with serum Tg levels <1 ng/mL during thyroid hormone suppression of TSH (785) will have a serum Tg level >2 ng/mL after rhTSH or thyroid hormone withdrawal at 12 months after initial therapy with surgery and RAI. In this patient population, one-third will have identification of persistent or recurrent disease and of increasing Tg levels, and the other two-thirds will remain free of clinical disease and will have stable or decreasing stimulated serum Tg levels over time (618,624). However, there may be a low likelihood of a rise in serum Tg to >2 ng/mL when the basal serum Tg is <0.1 ng/mL if a second-generation Tg immunoochemiluminometric assay (ICMA) with a functional sensitivity of 0.05 ng/mL is employed (788). There is good evidence that a Tg cutoff level above 2 ng/mL following rhTSH stimulation is highly sensitive in identifying patients with persistent tumor (785,789–794). However, the results of serum Tg measurements made on the same serum specimen differ among assay methods (316). Therefore, the Tg cutoff may differ significantly among medical centers and laboratories. Further, the clinical significance of minimally
detectable Tg levels is unclear, especially if only detected following TSH stimulation. However, receiver operating curves have shown that a Tg level on thyroid hormone around 0.2–0.3 ng/mL portends the best sensitivity and specificity for detecting persistent disease. In these patients, the trend in serum Tg over time will typically identify patients with clinically significant residual disease. A rising unstimulated or stimulated serum Tg indicates disease that is likely to become clinically apparent (618,795). Thyrogblobulin doubling time may have utility as a predictor of recurrence, analogous to the use of calcitonin doubling time for MTC (622,796).

[C7] Anti-Tg antibodies

The presence of anti-Tg antibodies, which occur in approximately 25% of thyroid cancer patients (797) and 10% of the general population (798), will falsely lower serum Tg determinations in immunometric assays (799). The use of recovery assays in this setting to detect significant interference is controversial (799,800). Serum anti-Tg antibody should be measured in conjunction with serum Tg assay by an immunometric method. Although assay standardization against the International Reference Preparation 65/93 has been recommended (608), wide-ranging variability in assay results and analytical sensitivity of the assay remains (801,802). Use of recovery methods for anti-Tg antibody may suffer variable interferences (608). Anti-Tg antibody may rise transiently postoperatively as an apparent immune reaction to the surgery itself and may also rise after ablation therapy (611). Anti-Tg antibodies should be measured in a different assay if the routine anti-Tg antibody assay is negative in a patient with surgically proven Hashimoto thyroiditis (775). It may be useful to measure anti-Tg antibodies shortly after thyroidectomy and prior to ablation because high levels may herald the likelihood of recurrence in patients without Hashimoto thyroiditis (801). Similarly, recurrent or progressive disease is suggested in those patients initially positive for anti-Tg antibodies who then become negative but subsequently have rising levels of anti-Tg antibodies. Falling levels of anti-Tg antibodies may indicate successful therapy (614,801). Thus, serial serum anti-Tg antibody quantification using the same methodology may serve as an imprecise surrogate marker of residual normal thyroid tissue, Hashimoto thyroiditis, or tumor (608,609,615). Following total thyroidectomy and RAI remnant ablation, anti-Tg antibodies usually disappear over a median of about 3 years in patients without evidence of persistent disease (611,615,616). Several studies demonstrate an increased risk of recurrence/persistent disease associated either with a new appearance of anti-Tg antibodies or rising titers (609–614). From a clinical perspective, anti-Tg antibody levels that are declining over time, especially for those with a rising Tg, remnant ablation or adjuvant therapy with 131I may be considered with a posttherapy WBS if neck US is negative. There is no evidence in these low-risk patients that a delayed treatment over the postoperative treatment may adversely affect the outcome.

A cohort of 80 consecutive patients with very low-risk PTMC who had undergone near-total thyroidectomy without postoperative RAI treatment were studied over 5 years (783). The rhTSH-stimulated serum Tg levels were ≤1 ng/mL in 45 patients (56%) and >1 ng/mL in 35 (44%) patients in whom rhTSH-stimulated Tg levels were as high as 25 ng/mL. The diagnostic WBS revealed uptake in the thyroid bed but showed no pathological uptake in any patient, and thyroid bed uptake correlated with the rhTSH-stimulated serum Tg levels (p < 0.0001). Neck ultrasonography identified lymph node metastases in both Tg-positive and Tg-negative patients. The authors concluded that for follow-up of this group of patients: (i) diagnostic WBS was ineffective at detecting metastases; (ii) neck ultrasonography as the main surveillance tool was highly sensitive in detecting lymph node metastases; and (iii) detectable rhTSH-stimulated serum Tg levels mainly depended upon the size of thyroid remnants, which suggests that serum Tg determination should be performed primarily on thyroid hormone therapy when using a sensitive Tg assay (functional sensitivity ≤0.2 ng/mL). In a series of 290 low-risk patients who had not undergone remnant ablation (545), serum Tg levels on LT4 became undetectable (<1 ng/mL) within 5–7 years in 95% of the cohort and was <0.1 ng/mL in 80% of a subset of these patients, using a sensitive assay to confirm the utility of Tg measurements on thyroid hormone treatment.

[C8] What is the role of serum Tg measurement in patients who have not undergone RAI remnant ablation?

■ RECOMMENDATION 64

Periodic serum Tg measurements on thyroid hormone therapy should be considered during follow-up of patients with DTC who have undergone less than total thyroidectomy and in patients who have had a total thyroidectomy but not RAI ablation. While specific cutoff levels of Tg that optimally distinguish normal residual thyroid tissue from persistent thyroid cancer are unknown, rising Tg values over time are suspicious for growing thyroid tissue or cancer.

(Strong recommendation, Low-quality evidence)
for routine follow-up. The frequency of follow-up is uncertain in patients who have not received RAI ablation and have sufficient residual thyroid tissue to produce measurable levels of serum Tg, the magnitude of which will depend upon the mass of residual tissue and the degree of TSH suppression. It appears reasonable to consider periodic measurements of Tg as surveillance for a trend in rising values.

[RECOMMENDATION 65]

Cervical ultrasonography

(A) Following surgery, cervical US to evaluate the thyroid bed and central and lateral cervical nodal compartments should be performed at 6–12 months and then periodically, depending on the patient’s risk for recurrent disease and Tg status.

(Strong recommendation, Moderate-quality evidence)

(B) If a positive result would change management, ultrasonographically suspicious lymph nodes ≥8–10 mm (see Recommendation 71) in the smallest diameter should be biopsied for cytology with Tg measurement in the needle washout fluid.

(Strong recommendation, Low-quality evidence)

(C) Suspicious lymph nodes less than 8–10 mm in smallest diameter may be followed without biopsy with consideration for FNA or intervention if there is growth or if the node threatens vital structures.

(Weak recommendation, Low-quality evidence)

(D) Low-risk patients who have had remnant ablation, negative cervical US, and a low serum Tg on thyroid hormone therapy in a sensitive assay (<0.2 ng/mL) or after TSH stimulation (Tg <1 ng/mL) can be followed primarily with clinical examination and Tg measurements on thyroid hormone replacement.

(Weak recommendation, Low-quality evidence)

Cervical ultrasonography is performed with a high-frequency probe (≥10 MHz) and is highly sensitive in the detection of cervical metastases in patients with DTC (290,783,807). These studies primarily evaluate patients with PTC, and the utility of neck US for monitoring patients with low-risk FTC is not well-established. Neck US should interrogate all lymph node compartments and the thyroid bed. Frequently, US does not distinguish thyroid bed recurrences from benign nodules (629,808). When an abnormality is found during the year after surgery in patients without any other suspicious findings, including low serum Tg on thyroid hormone therapy, follow-up may be performed with neck US. A correlation performed between US findings and pathology at surgery (292) has shown for lymph nodes >7 mm in the smallest diameter that a cystic appearance or hyperechoic punctuations in a context of thyroid cancer should be considered as malignant; lymph nodes with a hyperechoic hilum are reassuring; the type of vascularization (central: reassuring; peripheral: concerning) has a high sensitivity/specificity; a round shape, a hypoechoic appearance or the loss of the hyperechoic hilum by themselves does not justify a FNA biopsy (FNAB).

Interpretation of neck US should take into account all other clinical and biological data. In fact, the risk of recurrence is closely related to the initial lymph node status: most lymph node recurrences occur in already involved compartments; the risk increases with a higher number of N1 and a higher number of N1 with extracapsular extension (338) and with macroscopic rather than microscopic lymph node metastases (355,809).

In low- and intermediate-risk patients, the risk of lymph node recurrence is low (<2%) in patients with undetectable serum Tg and is much higher in those with detectable/elevated serum Tg. In fact, 1 g of neoplastic thyroid tissue will increase the serum Tg by ~1 ng/mL during LT4 treatment and by approximately 2–10 ng/mL following TSH stimulation (788,800). Neck US can detect N1 as small as 2–3 mm in diameter (in patients in whom serum Tg may be low or undetectable), but benefits of their early discovery (<8–10 mm) is not demonstrated.

FNAB for cytology and Tg measurement in the aspirate fluid is performed for suspicious lymph nodes ≥8–10 mm in their smallest diameter. US guidance may improve the results of FNAB, in particular for small lymph nodes and those located deep in the neck. However, FNAB cytology misses thyroid cancer in a significant proportion (up to 20%) of patients. The combination of cytology and serum Tg determination in the aspirate fluid increases sensitivity (303,810,811). In cases of lymph node metastases, the Tg concentration in the aspirate fluid is often elevated (>10 ng/mL), and concentrations above this level are highly suspicious (296,298,301). A Tg concentration in the aspirate fluid between 1 and 10 ng/mL is moderately suspicious for malignancy, and comparison of the Tg measurement in the aspirate fluid and the serum should be considered in these patients. Also, up to half of the FNAB performed for suspicious US findings are benign, demonstrating that selection of patients for FNAB needs to be improved (296,298,812). Nonsuspicious and small nodes (<8–10 mm in the smallest diameter) can be monitored with neck US.

[DIAGNOSTIC WHOLE-BODY RAI SCANS]

RECOMMENDATION 66

After the first posttreatment WBS performed following RAI remnant ablation or adjuvant therapy, low-risk and intermediate-risk patients (lower risk features) with an undetectable Tg on thyroid hormone with negative anti-Tg antibodies and a negative US (excellent response to therapy) do not require routine diagnostic WBS during follow-up.

(Strong recommendation, Moderate-quality evidence)

RECOMMENDATION 67

(A) Diagnostic WBS, either following thyroid hormone withdrawal or rhTSH, 6–12 months after adjuvant RAI therapy can be useful in the follow-up of patients with high or intermediate risk (higher risk features) of persistent disease (see risk stratification system, section [B19]) and should be done with 123I or low activity 131I.

(Strong recommendation, Low-quality evidence)

B) SPECT/CT RAI imaging is preferred over planar imaging in patients with uptake on planar imaging to better
Iodine 123 is preferred over 131I in these rare indications for even when neck US does not show any suspicious findings. with Tg antibodies, at risk of false-negative Tg measurement, zation of lower uptake in neck lymph nodes, and (iii) patients activity at the time of WBS) that may hamper the visuali-

informative postablation WBS because of large thyroid bed on posttherapy WBS, (ii) patients with poorly clinical settings: (i) patients with abnormal uptake outside the thyroid hormone and negative cervical US (583,785,813,814). 

residual tumor and have an undetectable serum Tg level on ally not necessary in low-risk patients who are clinically free of subsequent diagnostic WBSs have low sensitivity and are usu-

jects. It is used as a dosimetric and also as a diagnostic tool any uptake may be a cause of false negative SPECT-CT. MRI. Neoplastic lesions with low uptake of RAI or without further cross-sectional imaging studies such as contrast CT or MRI. Hybrid cameras combine a dual-head SPECT gamma camera with a CT scanner in one gantry. This allows direct superimposition of functional and anatomical imaging. The camera with a CT scanner in one gantry. This allows direct superimposition of functional and anatomical imaging. The

The sensitivity of 124I-PET for the detection of residual thyroid tissue and/or metastatic DTC was reported to be higher than that of a diagnostic 131I planar WBS (99% vs. 66%) (819–821). Iodine 124 PET/CT has not yet been compared with 131I SPECT/CT in a large series of patients with DTC. Furthermore, 124I is not yet widely available for clinical use and is primarily a research tool at this time.

[C12] 18FDG-PET scanning

- **RECOMMENDATION 68**

(A) 18FDG-PET scanning should be considered in high-risk DTC patients with elevated serum Tg (generally >10 ng/mL) with negative RAI imaging. 

(Strong recommendation, Moderate-quality evidence)

(B) 18FDG-PET scanning may also be considered as (i) a part of initial staging in poorly differentiated thyroid cancers and invasive Hürthle cell carcinomas, especially those with other evidence of disease on imaging or because of elevated serum Tg levels, (ii) a prognostic tool in patients with metastatic disease to identify lesions and patients at highest risk for rapid disease progression and disease-specific mortality, and (iii) an evaluation of post-treatment response following systemic or local therapy of metastatic or locally invasive disease.

(Weak recommendation, Low-quality evidence)

18FDG-PET/CT is primarily considered in high-risk DTC patients with elevated serum Tg (generally >10 ng/mL) with negative RAI imaging. In a meta-analysis of 25 studies that included 789 patients, the sensitivity of 18FDG-PET/CT was 83% (ranging from 50% to 100%) and the specificity was 84% (ranging from 42% to 100%) in non–131I-avid DTC (822). Factors influencing 18FDG-PET/CT sensitivity included tumor dedifferentiation, larger tumor burden, and to a lesser extent, TSH stimulation. 

18FDG-PET is more sensitive in patients with an aggressive histological subtype, including poorly differentiated, tall cell, and Hürthle cell thyroid cancer. 18FDG uptake on PET in metastatic DTC patients is a major negative predictive factor for response to RAI treatment and an independent prognostic factor for survival (823,824). It can also identify lesions with high 18FDG uptake (standardized uptake value) that may be more aggressive and should be targeted for therapy or close monitoring. It is complementary to 131I WBS, even in the presence of detectable 131I uptake in metastases, because 18FDG uptake may be present in neoplastic foci with no 131I uptake.

In patients with a TSH-stimulated Tg ≤10 ng/mL, the sensitivity of 18FDG is low, ranging from <10% to 30%. It is therefore recommended to consider 18FDG-PET only in DTC patients with a stimulated Tg level ≥10 ng/mL. Of course, this level needs to be adapted and lowered in case of aggressive pathological variant of thyroid cancer that may produce low amounts of serum Tg. Furthermore, in patients with undetectable Tg levels but with persistent Tg antibodies the level of serum Tg cannot be reliably assessed and 18FDG-PET may localize disease in some of these patients.

The sensitivity of 18FDG-PET scanning may be slightly increased with TSH stimulation. A multicentric prospective study on 63 patients showed an increase in the number of lesions
detected on the $^{18}$FDG-PET/CT performed after rhTSH stimulation compared to the $^{18}$FDG-PET/CT performed on thyroid hormone treatment and without TSH stimulation (825). However, the sensitivity for detecting patients with at least one tumor site was not improved by the rhTSH stimulation. In this study, the lesions found only by rhTSH-PET contributed adequately to an altered therapeutic plan in four patients (6%), and the clinical benefit of identifying these additional small foci remains to be proven. Its clinical benefit might be higher in patients with normal neck and chest CT scan and normal neck ultrasonography. A meta-analysis on seven studies including the previous study and comprising 168 patients confirmed these results and showed that $^{18}$FDG-PET/CT performed following TSH stimulation altered clinical management in only 9% of patients. Furthermore, false positives can be seen with $^{18}$FDG-PET imaging with or without TSH stimulation (825).

Results of $^{18}$FDG-PET/CT might alter the indications for $^{131}$I treatment or the decision for surgical removal of small tumor foci with $^{18}$FDG uptake. The frequency of false-positive lesions varies among series from 0% to 39%, and this high number justifies a FNAB with cytology and Tg measurement in the aspirate fluid in cases in which surgery is planned, based on $^{18}$FDG-PET results. The higher sensitivity of neck ultrasonography for the detection of small metastatic lymph nodes should be noted, with $^{18}$FDG-PET being more sensitive for some locations such as the retropharyngeal or the retro-clavicular regions (825).

To date, there is no evidence that TSH stimulation improves the prognostic value of $^{18}$FDG-PET imaging.

[C13] CT and MRI

**RECOMMENDATION 69**

(A) Cross-sectional imaging of the neck and upper chest (CT, MRI) with IV contrast should be considered (i) in the setting of bulky and widely distributed recurrent nodal disease where US may not completely delineate disease, (ii) in the assessment of possible invasive recurrent disease where potential aerodigestive tract invasion requires complete assessment, or (iii) when neck US is felt to be inadequately visualizing possible neck nodal disease (high Tg, negative neck US).

*(Strong recommendation, Moderate-quality evidence)*

(B) CT imaging of the chest without IV contrast (imaging pulmonary parenchyma) or with IV contrast (to include the mediastinum) should be considered in high-risk DTC patients with elevated serum Tg (generally >10 ng/mL) or rising Tg antibodies with or without negative RAI imaging.

*(Strong recommendation, Moderate-quality evidence)*

(C) Imaging of other organs including MRI brain, MR skeletal survey, and/or CT or MRI of the abdomen should be considered in high-risk DTC patients with elevated serum Tg (generally >10 ng/mL) and negative neck and chest imaging who have symptoms referable to those organs or who are being prepared for TSH-stimulated RAI therapy (withdrawal or rhTSH) and may be at risk for complications of tumor swelling.

*(Strong recommendation, Low-quality evidence)*

In patients with elevated or rising Tg or anti-Tg antibodies and NED on neck US or RAI imaging (if performed), CT imaging of the neck and chest should be considered. The frequency of positive anatomic imaging increases with higher serum Tg levels above 10 ng/mL. CT is the most frequently recommended first-line technique to search for lymph node metastases in patients with squamous cell carcinoma of the head and neck, and an injection of contrast medium is mandatory for the analysis of the neck and mediastinum (826). Radioiodine can be administered 4–8 weeks following the injection of contrast medium, because at that time a majority of the iodine contamination has disappeared in most patients (315). If there is a concern, a random urine iodine (and creatinine) prior to initiation of a LID and RAI testing or treatment can be measured to make sure the urine iodine is not high. Diagnostic CT scan may complement neck US for the detection of macrometastases in the central compartment, in the mediastinum, and behind the trachea (307–309), and it is the most sensitive tool for the detection of micrometastases in the lungs. Before revision surgery is contemplated, presumptive recurrent neck targets must be defined by high-resolution radiographic anatomic studies such as US or spiral axial CT scan to complement $^{18}$FDG-PET/CT or RAI imaging and must be carefully defined to allow for adequate preoperative mapping and definitive surgical localization. In addition to nodal assessment axial scanning, including CT scan with contrast has utility in the evaluation of locally recurrent invasive disease and relationships with vessels. Such patients may present with hoarseness, vocal cord paralysis on laryngeal exam, progressive dysphagia or mass fixation to surrounding structures, respiratory symptoms including stridor or hemoptysis, and lesions with rapid progression/enlargement. Such lesions are incompletely evaluated with US alone, and axial CT scanning with contrast medium is indicated.

The use of MRI has also been advocated for imaging the neck and the mediastinum. It is performed with and without injection of gadolinium chelate as a contrast medium and does not require any injection of iodine contrast medium. The performance of MRI for imaging the neck and mediastinum has not been directly compared with CT on large numbers of thyroid cancer patients (827–829). Compared to CT scan, it may better delineate any involvement of the aerodigestive tract (830,831). It is often used as second-line imaging technique in patients with demonstrated or suspicious lesions on CT scan in order to better delineate these lesions. In the lower part of the neck, movements of the aerodigestive axis during the procedure that may last several minutes will decrease the quality of images (414). Endoscopy of the trachea and esophagus, with or without ultrasonography, looking for evidence of intraluminal extension can also be helpful in cases of suspected aerodigestive tract invasion. MRI is less sensitive than CT scan for the detection of lung micronodules.

Finally, whether these imaging techniques (CT and MRI) should be performed for diagnostic purposes or whether an $^{18}$FDG-PET/CT scan should be performed as the first-line imaging procedure for diagnosis is still a matter of debate. In the past, CT scan with injection of contrast medium was more sensitive for the detection of lymph node metastases (832), but with modern PET/CT equipment, the CT scan of the PET/CT is as reliable as a CT scan used for radiology, and many lesions can be found on $^{18}$FDG-PET/CT scanning, even if no
injection of contrast medium has been performed (833,834). Delineation between lymph node metastases or local recurrence and vessels or the aerodigestive axis is often not well visualized on \( ^{18} \)FDG-PET/CT in the absence of contrast injection, and if necessary other imaging techniques (CT and MRI with contrast medium) may be performed especially for a preoperative work-up. As a result, most patients with extensive disease should be considered for \( ^{18} \)FDG-PET/CT and CT imaging with contrast, and some patients will also be considered for MRI.

This imaging strategy is applied in patients with elevated serum Tg (>5–10 ng/mL) and no other evidence of disease (neck and chest imaging), starting with a \( ^{18} \)FDG-PET/CT (822,833). In the past an empiric treatment was used in such patients, but recent studies have shown that \( ^{18} \)FDG-PET/CT imaging is more sensitive and should be performed as the first-line approach, with empiric RAI treatment being considered only for those patients with no detectable \( ^{18} \)FDG uptake (833).

[C14] Using ongoing risk stratification (response to therapy) to guide disease long-term surveillance and therapeutic management decisions

Ongoing risk stratification allows the clinician to continue to provide individualized management recommendations as the risk estimates evolve over time. While the specific details of how surveillance and therapeutic strategies should be modified over time as a function of response to therapy reclassification within each ATA risk category remains to be defined, we do endorse the following concepts (more details in Table 13).

Excellent response: An excellent response to therapy should lead to a decrease in the intensity and frequency of follow-up and the degree of TSH suppression (this change in management will be most apparent in ATA intermediate- and high-risk patients).

Biochemical incomplete response: If associated with stable or declining serum Tg values, a biochemical incomplete response should lead to continued observation with ongoing TSH suppression in most patients. Rising Tg or anti-Tg antibody values should prompt additional imaging and potentially additional therapies.

Structural incomplete response: A structural incomplete response may lead to additional treatments or ongoing observation depending on multiple clinico-pathologic factors including the size, location, rate of growth, RAI avidity, \( ^{18} \)FDG avidity, and specific pathology of the structural lesions.

Indeterminate response: An indeterminate response should lead to continued observation with appropriate serial imaging of the nonspecific lesions and serum Tg monitoring. Nonspecific findings that become suspicious over time or rising Tg or anti-Tg antibody levels can be further evaluated with additional imaging or biopsy.

[C15] What is the role of TSH suppression during thyroid hormone therapy in the long-term follow-up of DTC?

- **RECOMMENDATION 70**

(A) In patients with a structural incomplete response to therapy, the serum TSH should be maintained below 0.1 mU/L indefinitely in the absence of specific contraindications.

(B) In patients with a biochemical incomplete response to therapy, the serum TSH should be maintained between 0.1 and 0.5 mU/L, taking into account the initial ATA risk classification, Tg level, Tg trend over time, and risk of TSH suppression.

(C) In patients who have not undergone remnant ablation or adjuvant therapy who demonstrate an excellent or indeterminate response to therapy, especially those at low risk for recurrence, the serum TSH may be kept within the low reference range (0.5–2 mU/L).

(D) In patients with a biochemical incomplete response to therapy, especially those at low risk for recurrence, the serum TSH may be kept within the low reference range (0.5–2 mU/L).

(E) In patients who have not undergone remnant ablation or adjuvant therapy who demonstrate an excellent or indeterminate response to therapy with a normal neck US, and low or undetectable suppressed serum Tg, and Tg or anti-Tg antibodies that are not rising, the serum TSH can be allowed to rise to the low reference range (0.5–2 mU/L).

(F) In patients with high-risk disease but no detectable \( ^{18} \)FDG uptake, the serum TSH should be suppressed to 0.05 mU/L or less provided no ad-

A meta-analysis has suggested an association between thyroid hormone suppression therapy and reduction of major adverse clinical events (745). The appropriate degree of TSH suppression by thyroid hormone therapy is still unknown, especially in high-risk patients rendered free of disease. A constantly suppressed TSH (0.05 mU/L) was found in one study to be associated with a longer relapse-free survival than when serum TSH levels were always 1 mU/L or greater, and the degree of TSH suppression was an independent predictor of recurrence in multivariate analysis (749). Conversely, another large study found that stage disease, patient age, and \( ^{131} \)I therapy independently predicted disease progression, but the degree of TSH suppression did not (275). A third study showed that during LT₃ therapy the mean Tg levels were significantly higher when TSH levels were normal than when TSH levels were suppressed (<0.5 mU/L) but only in patients with local or distant recurrence (835). A fourth study of 2936 patients found that overall survival improved significantly when the TSH was suppressed to <0.1 mU/L in patients with NTCTCSG stage III or IV disease and to a range of 0.1 mU/L to about 0.5 mU/L in patients with NTCTCSG stage II disease; however, there was no incremental benefit from suppressing TSH to undetectable levels in stage II patients, and suppression of TSH was of no benefit in patients with stage I disease (671), and higher degrees of suppression to TSH of <0.03 mU/L provided no additional benefit (746). Another study found that a serum TSH
threshold of 2 mU/L differentiated best between patients free of disease and those with relapse or cancer-related mortality, which remained significant when age and tumor stage were included in a multivariate analysis (750). A prospective study showed that disease-free survival for low-risk patients without TSH suppression was not inferior to patients with TSH suppression (836). No prospective studies have been performed examining the risk of recurrence and death from thyroid cancer associated with varying serum TSH levels, based on the criteria outlined above in [C14] for the absence of tumor at 6–12 months post surgery and RAI ablation.

A recent observational study demonstrated increased risk of all-cause and cardiovascular mortality in DTC patients compared to a control population (837). The authors also showed that survival in the DTC patients was lower when the serum TSH was <0.02 mU/L, which is particularly relevant in patients with an excellent response to therapy in whom overtreatment should be avoided. An approach to balancing the risks of thyroxine suppression against the risks of tumor recurrence or progression has been presented in a recent review (747). This review helped to define patients at low, intermediate, and high risk of complications from TSH suppression therapy. Table 15 provides recommendations for TSH ranges based on response to thyroid cancer therapy weighed against risks of LT4 therapy, which is adapted from the review by Biondi and Cooper (747).

There are inadequate data to make a strong recommendation regarding the intensity and duration of TSH suppression in the biochemical incomplete response to therapy category. This category encompasses a variety of patients with low serum Tg levels (median nonstimulated Tg 3.6 ng/mL) having been initially classified as ATA low risk (16%–24%), ATA intermediate risk (47%–64%), or ATA high risk (18%–21%) (539,607). Furthermore, the risk of development of structurally identifiable disease within this cohort is not uniform but rather is related to the ongoing behavior of residual disease as reflected by both the magnitude of the Tg elevation and to the rate of rise of the serum Tg or anti-Tg antibodies. Based on weak data and expert opinion, we recommend a goal TSH of 0.1–0.5 mIU/L for the majority of patients with a biochemical

---

**Table 15. Thyrotropin Targets for Long-Term Thyroid Hormone Therapy**

<table>
<thead>
<tr>
<th>Increasing Risk of TSH Suppression</th>
<th>Excellent</th>
<th>Indeterminate</th>
<th>Biochemical Incomplete **</th>
<th>Structural Incomplete</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Known Risk</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Menopause</td>
<td>Yes</td>
<td>Yes</td>
<td>Moderate or Complete</td>
<td>Yes</td>
</tr>
<tr>
<td>Tachycardia</td>
<td>Yes</td>
<td>Yes</td>
<td>Complete Suppression</td>
<td>Yes</td>
</tr>
<tr>
<td>Osteopenia</td>
<td>Yes</td>
<td>Yes</td>
<td>Complete Suppression</td>
<td>Yes</td>
</tr>
<tr>
<td>Age &gt; 60</td>
<td>Yes</td>
<td>Yes</td>
<td>Complete Suppression</td>
<td>Yes</td>
</tr>
<tr>
<td>Osteoporosis</td>
<td>Yes</td>
<td>Yes</td>
<td>Complete Supression</td>
<td>Yes</td>
</tr>
<tr>
<td>Atrial Fibrillation</td>
<td>Yes</td>
<td>Yes</td>
<td>Complete Supression</td>
<td>Yes</td>
</tr>
</tbody>
</table>

---

*0.5 mU/L represents the lower limit of the reference range for the TSH assay which can be 0.3–0.5 mU/L depending on the specific assay

**TSH target for patients with a biochemical Incomplete response can be quite different based on original ATA risk, Tg level, Tg trend over time and risk of TSH suppression

- No suppression. TSH target 0.5–2.0 mU/L
- Mild suppression. TSH target 0.1–0.5 mU/L
- Moderate or Complete suppression. TSH target <0.1 mU/L
incomplete response, recognizing that less intense TSH suppression (0.5–2.0 mIU/L) may be appropriate for ATA low-risk patients with stable nonstimulated Tg values near the threshold for excellent response (e.g., nonstimulated Tg values in the 1–2 ng/mL range), while more intense TSH suppression (<0.1 mIU/L) may be desired in the setting of more elevated or rapidly rising Tg values.

[C16] What is the most appropriate management of DTC patients with metastatic disease?

Metastases may be discovered at the time of initial disease staging or may be identified during longitudinal follow-up. If metastases are found following initial therapy, some patients may subsequently experience a reduction in tumor burden with additional treatments that may offer a survival or palliative benefit (840–844). The preferred hierarchy of treatment for metastatic disease (in order) is surgical excision of locoregional disease in potentially curable patients, 131I therapy for RAI-responsive disease, external beam radiation therapy or other directed treatment modalities such as thermal ablation, TSH-suppressive thyroid hormone therapy for patients with stable or slowly progressive asymptomatic disease, and systemic therapy with kinase inhibitors (preferably by use of FDA-approved drugs or participation in clinical trials), especially for patients with significantly progressive macroscopic refractory disease. Clinical trials or kinase inhibitor therapy may be tried before external beam radiation therapy in special circumstances, in part because of the morbidity of external beam radiation and its relative lack of efficacy. However, localized treatments with thermal (radiofrequency or cryo-) ablation (845), ethanol ablation (846), or chemo-embolization (847) may be beneficial in patients with a single or a few metastases and in those with metastases at high risk of local complications; the treatments should be performed in such patients before the initiation of any systemic treatment. These modalities may control treated metastases, may avoid local complications, and may delay initiation of systemic treatment. Additionally, surgical therapy in selected incurable patients is important to prevent complications in targeted areas, such as the CNS and central neck compartment. Conversely, conservative intervention with TSH-suppressive thyroid hormone therapy may be appropriate for selected patients with stable asymptomatic local metastatic disease and most patients with stable asymptomatic non-CNS distant metastatic disease.

[C17] What is the optimal directed approach to patients with suspected structural neck recurrence?

- RECOMMENDATION 71

Therapeutic compartmental central and/or lateral neck dissection in a previously operated compartment, sparing uninvolved vital structures, should be performed for patients with biopsy-proven persistent or recurrent disease for central neck nodes ≥8 mm and lateral neck nodes ≥10 mm in the smallest dimension that can be localized on anatomic imaging.

(Strong recommendation, Moderate-quality evidence)

Persistent or recurrent nodal disease may result in local invasion and is the source of considerable patient and physician anxiety (848). However, several observational studies suggest that low-volume recurrent nodal disease can be indolent and can be managed through active surveillance, although not all lesions in these series are documented as malignant (629,849). Bulky or invasive recurrent disease is best treated surgically (319,850–853).

The judgment to offer surgery for recurrent nodal disease in the neck is made with equipoise in two opposing decision elements: (i) the risks of revision surgery (which are typically higher than primary surgery due to scarring from previous surgery (854) balanced with (ii) the fact that surgical resection generally represents the optimal treatment of macroscopic gross nodal disease over other treatment options. An important element in this decision-making process is the availability of surgical expertise specifically in the performance of revision thyroid cancer nodal surgery, which is a discrete surgical skill set. The decision to treat cervical nodal recurrence surgically should be made with an appreciation of distant disease presence and progression but may be undertaken even in the setting of known distant metastasis for palliation of symptoms and prevention of aerodigestive tract obstruction. The decision for treatment and surgery specifically is best derived through collaborative team approach involving surgery, endocrinology, and importantly the patient and family (855). Therefore, cytologic confirmation of disease can be deferred if the findings of the FNA will not lead to additional evaluation or treatment. While we generally recommend cytologic confirmation of abnormal radiographic findings prior to surgical resection, we recognize that this may not be necessary (or possible) in every case (e.g., radiographic findings with a very high likelihood of malignancy, or the specific location of the lymph node makes it difficult/impossible to biopsy).

[C18] Nodal size threshold

Surgery is considered with the recognition of clinically apparent, macroscopic nodal disease through radiographic analysis including US (Table 7) and/or axial (CT) scanning rather than through isolated Tg elevation (309,335,856). Given the risks of revision nodal surgery, a clearly defined preoperative radiographic target is mandatory. The risks of surgery relate in part to the exact location of the target node(s) and whether the compartment in question has been previously dissected such as recurrent central neck nodes after primary thyroidectomy. This target must be defined by high-resolution radiographic anatomic studies such as US or spiral CT scan with contrast, as a complement to 18FDG-PET/CT or RAI-SPECT/CT when performed, to allow for adequate preoperative mapping and definitive surgical localization (309,856). Ultrasound-guided FNA for cytology with Tg measurement in the aspiration sample can be performed in the setting of radiographically suspicious nodal recurrence keeping in mind that Tg rinsing may be positive with thyroid bed persistent benign thyroid remnant tissue if the patient has not been treated with RAI. Charcoal tattooing under US guidance may help the surgeon to localize the lymph node to be removed during surgery (344).

Malignant central neck nodes ≥8 mm and lateral neck nodes ≥10 mm in the smallest dimension that have undergone FNAB and can be localized on anatomic imaging (US with or without axial CT) can be considered surgical targets.
Short-axis nodal diameter measurement is optimally employed in surgical decision-making for nodal malignancy. Smaller lesions are probably best managed with active surveillance (observation) with serial cross-sectional imaging, reserving FNA and subsequent intervention for documented structural disease progression. However, multiple factors in addition to size should be taken into account when considering surgical options, including proximity of given malignant nodes to adjacent vital structures and the functional status of the vocal cords. Patient comorbidities, motivation, and emotional concerns should also be taken into account along with primary tumor factors (high-grade histology, Tg doubling time, RAI avidity, 18FDG-PET avidity, and presence of molecular markers associated with aggressive behavior). Through thorough patient and multidisciplinary collaborative discussions, metastatic nodes >8–10 mm can be carefully observed in properly selected patients with serial clinical and radiographic follow-up, with surgery being offered if they progress during follow-up and conservative follow-up being maintained if they are stable over time.

**[C19] Extent of nodal surgery**

Because of the increased risk of recurrence with focal “berry-picking” techniques, compartmental surgery is recommended (860,861). Planned compartmental dissection should be adjusted and be more limited depending on the surgeon’s judgment of procedural safety as it relates to scarring/distortion of anatomy (from prior surgery and/or past radiation therapy) and the perception of impending complications. Typical revision lateral neck dissection involves levels II, III, and IV, while revision central neck dissection includes at least one paratracheal region with prelaryngeal and pretracheal subcompartments. Bilateral central neck dissection is offered only when dictated by disease distribution because of the risks of bilateral nerve injury and permanent hypoparathyroidism.

Basal Tg decreases by 60%–90% after compartmental dissection for recurrent nodal disease in modern series, but only 30%–50% of patients have unmeasurable basal Tg after such surgery, and it is difficult to predict who will respond to surgery with Tg reduction (627,628,634,858,862–876). However, most series suggest surgery results in a high clearance rate of structural disease in over 80% of patients (859,875).

**[C20] Ethanol injection**

Percutaneous ethanol injection for patients with metastatic lymph nodes is gaining interest as a nonsurgical directed therapy for patients with recurrent DTC. Most of the studies limited PEI to patients who had undergone previous neck dissections and RAI treatment, those who had FNA-proven DTC in the lymph node, and those with no known distant metastases.

One of the first studies examining the effectiveness of local metastatic lymph node control by PEI treated 14 patients with 29 lymph nodes (846). Twelve of the 14 patients had good loco-regional control in this study with short-term follow-up (mean 18 months). The largest study to date treated 63 patients with 109 metastatic lymph nodes between the years 2004 and 2009 (878). Ninety-two lymph nodes (84%) were successfully ablated in this retrospective study with a mean follow-up of 38 months, and most required one to three treatment sessions. Minor complications included brief discomfort at the PEI site, and there were no major complications.

A recent study retrospectively reviewed 25 patients who had 37 lymph nodes ablated between the years 1994 and 2012, with a relatively long follow-up of a mean of 65 months (879). All lymph nodes were successfully ablated in one to five treatment sessions by lack of flow on US. Most of the lymph nodes decreased in size and 46% completely disappeared. Serum Tg levels were reduced in most patients and brought into an acceptable range (<2.4 ng/mL) in 82% of patients with negative anti-Tg antibodies. There were no serious or long-term complications. Another recent study also demonstrated safety and efficacy of PEI in 21 patients with 41 metastatic lymph nodes (880). These investigators treated patients with only one session, and 24% of patients had a recurrence at the site of the injection.

Limitations of many of the studies included small numbers of patients, relatively short-term follow-up, and many patients with small lymph nodes (<5–8 mm).

A general consensus from studies and reviews is that PEI should be considered in patients who are poor surgical candidates. Many patients will likely need more than one treatment session and lymph nodes >2 cm may be difficult to treat with PEI. Focal PEI treatment does represent a nonsurgical form of berry picking. Formal neck compartmental dissection is still the first-line therapy in DTC patients with clinically apparent or progressive lymph node metastases. When deciding for the optimal strategy of care for a patient’s lymph node metastases, previous treatment modalities should also be taken into consideration.

**[C21] Radiofrequency or laser ablation**

The use of radiofrequency ablation (RFA) with local anesthesia in the treatment of recurrent thyroid cancer has been associated with a mean volume reduction that ranges between approximately 55% and 95% (881,882) and complete disappearance of the metastatic foci in 40%–60% of the cases (845,882,883). As with alcohol ablation, multiple treatment sessions are often required. Complications include discomfort, pain, skin burn, and changes in the voice (884). Similar to alcohol ablation techniques, it appears that RFA may be most useful in high-risk surgical patients or in patients refusing additional surgery, rather than as a standard alternative to surgical resection of metastatic disease (883–885). More recently, preliminary findings using US-guided laser ablation for treatment of cervical lymph node metastases have been reported (886).

**[C22] Other therapeutic options**

Empiric RAI therapy for structurally identifiable disease that is not RAI avid by diagnostic scanning is very unlikely to have a significant tumoricidal effect and is therefore not generally recommended (887). Stereotactic radiotherapy (SBRT) can be successfully used to treat isolated metastatic disease foci, but it has no role in most patients with resectable lymph node metastases. EBRT using modern techniques such as intensity modulated radiotherapy and stereotactic radiation, is considered for loco-regional recurrence that is not surgically resectable or with extranodal extension or involvement of soft tissues, particularly in patients with no evidence of
distant disease. Efficacy has been suggested only in retrospective studies on limited numbers of patients (888,889). Likewise, systemic therapies (such as cytotoxic chemotherapy or kinase inhibitors) for loco-regional disease are considered only after all surgical and radiation therapy options have been exhausted.

[C23] What is the surgical management of aerodigestive invasion?

■ RECOMMENDATION 72

When technically feasible, surgery for aerodigestive invasive disease is recommended in combination with RAI and/or EBRT.

(Strong Recommendation, Moderate-quality evidence)

For tumors that invade the upper aerodigestive tract, surgery combined with additional therapy such as $^{131}$I and/or external beam radiation therapy is generally advised (890,891). Patient outcome is related to complete resection of all gross disease with the preservation of function, with techniques ranging from shaving a tumor off the trachea or esophagus for superficial invasion, to more aggressive techniques when the trachea is more deeply invaded (e.g., direct intraluminal invasion), including tracheal resection and Anastomosis or laryngopharyngoesophagectomy (892–894). Surgical decision-making can be complex and must balance oncologic surgical completeness with preservation of upper aerodigestive tract head and neck function. In some circumstances such surgery represents a possible attempt for cure, and in other circumstances it offers significant regional neck palliation in patients with distant metastasis with impending asphyxiation or significant hemoptysis (414,891).

[C24] How should RAI therapy be considered for loco-regional or distant metastatic disease?

For regional nodal metastases discovered on diagnostic WBS, RAI may be employed in patients with low-volume disease or in combination with surgery, although surgery is typically preferred in the presence of bulky disease or disease amenable to surgery. Radiiodine is also used adjunctively following surgery for regional nodal disease or aerodigestive invasion if residual RAI-avid disease is present or suspected. Significant variation exists nationally in the United States in regard to RAI use, irrespective of the degree of disease or risk of recurrence (895,896). However, there are no randomized, controlled clinical trials demonstrating better patient outcomes after RAI therapy. One retrospective analysis indicated that a delay in RAI therapy of 6 months or more was associated with disease progression and reduced survival (897). In one study of 45 patients with persistent serum Tg elevation after reoperation for loco-regional recurrence, adjuvant RAI therapy demonstrated no benefit (898).

[C25] Administered activity of $^{131}$I for loco-regional or metastatic disease

■ RECOMMENDATION 73

(A) Although there are theoretical advantages to dosimetric approaches to the treatment of loco-regional or metastatic disease, no recommendation can be made about the superiority of one method of RAI administration over another (empiric high activity versus blood and/or body dosimetry versus lesional dosimetry).

(No recommendation, Insufficient evidence)

(B) Empirically administered amounts of $^{131}$I exceeding 150 mCi that often potentially exceed the maximum tolerable tissue dose should be avoided in patients over age 70 years.

(Strong recommendation, Moderate-quality evidence)

Despite the apparent effectiveness of $^{131}$I therapy in many patients, the optimal therapeutic activity remains uncertain and controversial (895,899). There are three approaches to $^{131}$I therapy: empiric fixed amounts, therapy determined by the upper limit of blood and body dosimetry (900–907), and quantitative tumor or lesional dosimetry (908,909). Dosimetric methods are often reserved for patients with distant metastases or unusual situations such as renal insufficiency (910,911), children (912,913), the elderly, and those with extensive pulmonary metastases (914). Comparison of outcomes among these methods from published series is difficult (895,897,900). No prospective randomized trial to address the optimal therapeutic approach has been published. One retrospective study concluded that patients with loco-regional disease were more likely to respond after dosimetric therapy than after empiric treatment (915). Another study demonstrated improved efficacy of administration of dosimetric maximal activity after failure of empiric dosage (916). Arguments in favor of higher activities cite a positive relationship between the total $^{131}$I uptake per tumor mass and outcome (908), while others have not confirmed this relationship (917). In the future, the use of $^{123}$I or $^{131}$I with modern SPECT/CT or $^{124}$I PET-based dosimetry may facilitate whole-body and lesional dosimetry (918–920). In certain settings (e.g., in the patients with radiation-induced thyroid cancer seen after the Chernobyl accident), RAI treatment may be associated with good outcomes even in high-risk patients with metastatic disease (921).

The efficacy of RAI therapy is related to the mean radiation dose delivered to neoplastic foci and also to the radiosensitivity of tumor tissue (922). The radiosensitivity is higher in patients who are younger, with small metastases from well-differentiated papillary or follicular carcinoma and with uptake of RAI but no or low $^{18}$FDG uptake.

The maximum tolerated radiation absorbed dose (MTRD), commonly defined as 200 rads (cGy) to the blood, is potentially exceeded in a significant number of patients undergoing empiric treatment with various amounts of $^{131}$I. In one study (923) 1%–22% of patients treated with $^{131}$I according to dosimetry calculations would have theoretically exceeded the MTRD had they been empirically treated with 100–300 mCi of $^{131}$I. Another study (924) found that an empirically administered $^{131}$I activity of 200 mCi would exceed the MTRD in 8%–15% of patients younger than age 70 and 22%–38% of patients aged 70 years and older. Administering 250 mCi empirically would have exceeded the MTRD in 22% of patients younger than 70 and 50% of patients 70 and older. These estimates imply the need for caution in administering empiric activities higher than 100–150 mCi in certain populations such as elderly patients and patients with renal insufficiency.
**RECOMMENDATION 74**

There are currently insufficient outcome data to recommend rhTSH-mediated therapy for all patients with distant metastatic disease being treated with $^{131}$I.

*(No Recommendation, Insufficient evidence)*

**RECOMMENDATION 75**

Recombinant human TSH–mediated therapy may be indicated in selected patients with underlying comorbidities making iatrogenic hypothyroidism potentially risky, in patients with pituitary disease whose serum TSH cannot be raised, or in patients in whom a delay in therapy might be deleterious. Such patients should be given the same or higher activity that would have been given had they been prepared with hypothyroidism or a dosimetrically determined activity.

*(Strong Recommendation, Low-quality evidence)*

No randomized trial comparing thyroid hormone withdrawal to rhTSH-mediated therapy for treatment of distant metastatic disease has been reported, but there is a growing body of nonrandomized studies exploring the use of rhTSH to prepare patients for therapy of metastatic disease (713,925–934). One small comparative study showed that the radiation dose to metastatic foci is lower with rhTSH than that following withdrawal (935). Many of these case reports and series report disease stabilization or improvement in some patients following rhTSH-mediated $^{131}$I therapy, but whether the efficacy of this preparation is comparable to thyroid hormone withdrawal is unknown. Extreme or prolonged elevations of TSH from either thyroid hormone withdrawal or rhTSH may acutely stimulate tumor growth and mass of metastases (930,936–938). With metastatic deposits in the brain or in close relation to the spinal cord or the superior vena cava, such swelling may severely compromise neurologic function or produce a superior vena cava syndrome, respectively. When distant metastases are evident, MRI of the brain and spine is recommended to detect the presence of critical metastases prior to treatment (see sections [C28] and [C39]). When metastases are detected, institution of temporary high-dose corticosteroid therapy is recommended for trying to limit the risk of acute tumor swelling and compromised function. Dexamethasone has been employed in doses of 2–4 mg every 8 hours starting 6–12 hours prior to rhTSH and RAI dosing or after 10–12 days of thyroid hormone withdrawal, with the steroids continued in a tapering dosage schedule for 1 week post therapy, for 48–72 hours after rhTSH administration, or for 72 hours after re-institution of thyroxine therapy when thyroid hormone withdrawal was employed (927). In patients with these critical metastases, consideration should be given to preparation with either a reduced dose of rhTSH or to attenuating the degree and duration of endogenous TSH elevation after thyroid hormone withdrawal while monitoring serum TSH levels. This can be achieved by the temporary addition of LT$_3$ therapy to thyroxine replacement. Satisfactory RAI treatment with either empiric or dosimetric activities should be feasible after achieving TSH levels of 30–50 mU/L. When thyroid hormone withdrawal has been employed, LT$_4$ therapy should recommence once the dose of RAI is administered in order to reduce the duration of TSH elevation.

*[C26] Use of rhTSH (Thyrogen) to prepare patients for $^{131}$I therapy for loco-regional or metastatic disease*

**RECOMMENDATION 76**

Since there are no outcome data that demonstrate a better outcome of patients treated with lithium as an adjunct to $^{131}$I therapy, the data are insufficient to recommend lithium therapy.

*(No recommendation, Insufficient evidence)*

Lithium inhibits iodine release from the thyroid without impairing iodine uptake, thus enhancing $^{131}$I retention in normal thyroid and tumor cells (939). One study (940) found that lithium increased the estimated $^{131}$I radiation dose in metastatic tumors on average by more than 2-fold, but primarily in those tumors that rapidly cleared iodine. On the other hand, a different study was unable to document any clinical advantage of lithium therapy on outcome in patients with metastatic disease, despite an increase in RAI uptake in tumor deposits (941). Lithium use may be associated with adverse events and needs to be precisely managed.

*[C27] Use of lithium in $^{131}$I therapy*

**RECOMMENDATION 78**

How should distant metastatic disease to various organs be treated?

The overall approach to treatment of distant metastatic thyroid cancer is based upon the following observations and oncologic principles:

1. Morbidity and mortality are increased in patients with distant metastases, but individual prognosis depends upon factors including histology of the primary tumor, distribution and number of sites of metastasis (e.g., brain, bone, lung), tumor burden, age at diagnosis of metastases, and $^{18}$FDG and RAI avidity (842,933,942–948).
2. Improved survival is associated with responsiveness to directed therapy (surgery, EBRT, thermal ablation, etc.) and/or RAI (842,933,942–948).
3. In the absence of demonstrated survival benefit, certain interventions can provide significant palliation or reduce morbidity (847,949–951).
4. Treatment of a specific metastatic area must be considered in light of the patient’s performance status and other sites of disease; for example, 5%–20% of patients with distant metastases die from progressive cervical disease (948,952).
5. Longitudinal re-evaluation of patient status and continuing reassessment of potential benefit and risk of intervention are required.
6. In the setting of overall poor anticipated outcome of patients with radiographically evident or symptomatic metastases that do not respond to RAI, the complexity of multidisciplinary treatment considerations and the availability of prospective clinical trials should encourage the clinician to refer such patients to tertiary centers with particular expertise.
7. Mutation profiling of metastatic tumor (to detect abnormalities in genes such as **BRAF**, **TERT**, **RAS**, or **PAX8/
RECOMMENDATION 78

AA Pulmonary micrometastases should be treated with RAI therapy and RAI therapy should be repeated every 6–12 months as long as disease continues to concentrate RAI and respond clinically because the highest rates of complete remission are reported in these subgroups.

(Strong recommendation, Moderate-quality evidence)

BB The selection of RAI activity to administer for pulmonary micrometastases can be empiric (100–200 mCi, or 100–150 mCi for patients ≥70 years old) or estimated by dosimetry to limit whole-body retention to 80 mCi at 48 hours and 200 cGy to the bone marrow.

(Strong recommendation, Moderate-quality evidence)

RECOMMENDATION 79

Radioiodine-avid macronodular metastases may be treated with RAI and treatment may be repeated when objective benefit is demonstrated (decrease in the size of the lesions, decreasing Tg), but complete remission is not common and survival remains poor. The selection of RAI activity to administer can be made empirically (100–200 mCi) or by lesional dosimetry or whole-body dosimetry if available in order to limit whole-body retention to 80 mCi at 48 hours and 200 cGy to the bone marrow.

(Weak recommendation, Low-quality evidence)

In the management of the patient with pulmonary metastases, key criteria for therapeutic decisions include (i) size of metastatic lesions (macronodular typically detected by chest radiography, micronodular typically detected by CT, lesions beneath the resolution of CT); (ii) avidity for RAI and, if applicable, response to prior RAI therapy; and (iii) stability (or lack thereof) of metastatic lesions. Pulmonary pneumonitis and fibrosis are rare complications of high-dose RAI treatment. Dosimetric approaches to therapy with a limit of 80 mCi whole-body retention at 48 hours and 200 cGy to the bone marrow should be considered in patients with diffuse 131I pulmonary uptake (955). If pulmonary fibrosis is suspected, then appropriate periodic pulmonary function testing and consultation should be obtained. The presence of pulmonary fibrosis may limit the ability to further treat metastatic disease with RAI (956).

Patients with pulmonary micrometastases (<2 mm, generally not seen on anatomic imaging) that are RAI avid have the highest rates of complete remission after treatment with RAI (942,947,957,958). These patients should be treated with RAI therapy repeatedly every 6–12 months as long as disease continues to concentrate RAI and respond clinically.

A precise definition of “responding clinically” is not feasible given the wide variation in disease presentation and response to therapy. A meaningful response to RAI treatment is generally associated with a significant reduction in serum Tg and/or in the size or rate of growth of metastases or structurally apparent disease. In contrast, a reduction in serum Tg and in RAI uptake with no concomitant decrease or with an increase in tumor size suggests refractoriness to RAI therapy.

In the presence of widespread metastases, especially when in bone, additional RAI may temporarily stabilize progression, but it is unlikely to result in cure. The risks of bone marrow suppression or pulmonary fibrosis should generate caution when repeated doses of RAI are being considered. Absolute neutrophil count and platelet counts are the usual markers of bone marrow suppression, and pulmonary function testing including diffusing capacity of the lungs for carbon monoxide can be markers of pulmonary toxicity. Other approaches (see sections [C36–C41]) should be considered once maximal cumulative tolerable radiation doses have been administered.

Macronodular pulmonary metastases may also be treated with RAI if demonstrated to be iodine avid. How many doses of RAI to give and how often to give RAI is a decision that must be individualized based on the disease response to treatment, age of the patient, and the presence or absence of other metastatic lesions (942,947). The presence of significant side effects including bone marrow suppression and salivary gland damage may temper enthusiasm for additional RAI therapy (959), and risk of second malignancies after RAI treatment remains controversial (766,960,961). Patients with solitary pulmonary DTC metastases may be considered for surgical resection, although the potential benefit weighed against the risk of surgery is unclear.

The likelihood of significant long-term benefit of 131I treatment in patients with elevated Tg and negative diagnostic RAI scans is very low. (962,963). While some reduction in serum Tg may be observed after such empiric therapy, one analysis concluded that there was no good evidence either for or against such treatment (964). In one small retrospective series of patients with structural disease but negative diagnostic 131I WBSs, additional RAI therapy was associated with stability of disease in 44%, but progression of structural disease occurred in 56% of the patients (887).

RAI therapy of iodine-avid bone metastases has been associated with improved survival and should be employed, although RAI is rarely curative.

(Strong recommendation, Moderate-quality evidence)

BB The RAI activity administered can be given empirically (100–200 mCi) or determined by dosimetry.

(Weak recommendation, Low-quality evidence)

RAI therapy for patients with bone metastases is rarely curative, but some patients with RAI-avid bone metastases may benefit from this therapy (842,947). A dosimetrically
determined administered dose of RAI may be beneficial for patients with bone metastases (908); although this is not proven in controlled studies. Patients undergoing RAI therapy for bone metastases should also be considered for directed therapy of bone metastases that are visible on anatomical imaging (section [C38]). This may include surgery, external beam radiation therapy, and other focal treatment modalities. These patients should also be considered for systemic therapy with bone-directed agents (section [C47]).

[C31] When should empiric RAI therapy be considered for Tg-positive, RAI diagnostic scan–negative patients?

RECOMMENDATION 80
In the absence of structurally evident disease, patients with stimulated serum Tg <10 ng/mL with thyroid hormone withdrawal or <5 ng/mL with rTSH (indeterminate response) can be followed without empiric RAI therapy on continued thyroid hormone therapy alone, reserving additional therapies for those with rising serum Tg levels over time or other evidence of structural disease progression.

(Weak recommendation, Low-quality evidence)

RECOMMENDATION 81
Empiric (100–200 mCi) or dosimetrically determined RAI therapy may be considered in patients with more significantly elevated serum Tg levels (see Recommendation 80), rapidly rising serum Tg levels, or rising anti-Tg antibody levels, in whom imaging (anatomic neck/chest imaging and/or 18FDG-PET/CT) has failed to reveal a tumor source that is amenable to directed therapy. The risk of high cumulative administered activities of RAI must be balanced against uncertain long-term benefits. If empiric RAI therapy is given and the posttherapy scan is negative, the patient should be considered to have RAI-refractory disease and no further RAI therapy should be administered.

(Weak recommendation, Low-quality evidence)

RECOMMENDATION 82
If persistent nonresectable disease is localized after an empiric dose of RAI, and there is objective evidence of significant tumor reduction, then consideration can be made for RAI therapy to be repeated until the tumor has been eradicated or the tumor no longer responds to treatment. The risk of repeated therapeutic doses of RAI must be balanced against uncertain long-term benefits.

(Weak recommendation, Low-quality evidence)

Factors to consider when selecting patients for empiric RAI therapy include the level of serum Tg elevation and the results of 18FDG-PET scanning, if performed. Since tumors that are 18FDG-PET positive generally do not concentrate RAI (965), RAI therapy is much less likely to be efficacious (965,966) and it is unlikely to alter the poorer outcome in such patients (823). Because of this, it is reasonable to perform 18FDG-PET/CT scanning prior to consideration of empiric RAI therapy (967).

The cutoff value of serum Tg above which a patient should be treated with an empiric dose of RAI is not clear. Most studies have reported primarily on patients with Tg levels after T4 withdrawal of 10 ng/mL or higher; it has been suggested that a corresponding level after rTSH stimulation would be 5 ng/mL (321,782,962,968,969). Patients with a suppressed (624) or stimulated (970) serum Tg of 5 ng/mL or higher are unlikely to demonstrate a decline without therapy, and they have higher rates of subsequent structural recurrence than those with lower serum Tg levels (970). In addition, a rising serum Tg indicates disease that is likely to become clinically apparent, particularly if it is rapidly rising (622,971,972).

If serum Tg levels suggest residual or recurrent disease, but diagnostic RAI WBS imaging is negative and structural imaging does not reveal disease that is amenable to directed therapy (surgical, thermal ablation, EBRT, alcohol ablation, see section [C17 and C38]), then empiric therapy with RAI (100–200 mCi) or dosimetrically determined RAI activities can be considered for two purposes: (i) to aid in disease localization, and/or (ii) as therapy for nonsurgical disease. This approach may identify the location of persistent disease in approximately 50% of patients (968,973,974), although the reported range of success is wide. From a therapeutic perspective, over half of patients experience a fall in serum Tg after empiric RAI therapy in patients with negative diagnostic WBS (963,969,975,976); however, there is no evidence for improved survival with empiric therapy in this setting (782,962,968). Further, there is evidence that Tg levels may decline without specific therapy in a significant proportion of patients with Tg levels <10 ng/mL (539,618–620,624,782,970–972,975,977–979). The most compelling evidence for benefit from empiric RAI therapy is for pulmonary metastases, which are not amenable to surgical management or EBRT (782,844,980).

[C32] What is the management of complications of RAI therapy?

RECOMMENDATION 83
The evidence is insufficient to recommend for or against the routine use of measures to prevent salivary gland damage after RAI therapy.

(No recommendation, Low-quality evidence)

RECOMMENDATION 84
Patients with xerostomia are at increased risk of dental caries and should discuss preventive strategies with their dental/oral health professional.

(Weak recommendation, Low-quality evidence)

RECOMMENDATION 85
Surgical correction should be considered for nasolacrimal outflow obstruction, which often presents as excessive tearing (epiphora) but also predisposes to infection.

(Strong recommendation, Low-quality evidence)

While RAI appears to be a reasonably safe therapy, it is associated with a cumulative dose-related low risk of early- and late-onset complications such as salivary gland damage, dental caries (981), nasolacrimal duct obstruction (982), and secondary malignancies (762,763,961,983,984), and it may likely contribute to long-term dysphagia (985). Therefore, it is important to ensure that the benefits of RAI therapy, especially repeated courses, outweigh the potential risks. There is probably no dose...
of RAI that is completely safe, nor is there any maximum cumulative dose that could not be used in selected situations. However, with higher individual and cumulative doses there are increased risks of side effects as discussed previously.

For acute transient loss of taste or change in taste and sialadenitis, recommended measures to prevent damage to the salivary glands have included hydration, sour candies, amifostine, and cholinergic agents (986), but evidence is insufficient to recommend for or against these modalities. One study suggested sour candy may actually increase salivary gland damage when given within 1 hour of RAI therapy, as compared to its use until 24 hours post therapy (987). Another study showed that the use of lemon slices within 20 min of 124I administration resulted in increased radiation absorbed dose to the salivary glands (988). A different study suggested that early use and multiple administered doses of lemon juice transiently decreased radiation exposure to the parotid glands (989), so the exact role and details of use of saliagettes to prevent salivary gland damage remains uncertain. Patients with painful sialadenitis may receive pain relief from local application of ice. For chronic salivary gland complications, such as dry mouth and dental caries, cholinergic agents may increase salivary flow (986). Interventional sialendoscopy has been shown in a number of small studies to be an effective treatment in patients with RAI-induced sialadenitis that is unresponsive to medical therapy (990–992).

[C34] What other testing should patients receiving RAI therapy undergo?

- **RECOMMENDATION 86**

Although patients should be counseled on the risks of second primary malignancies with RAI treatment for DTC, the absolute increase in risk of developing a second primary malignancy attributable to RAI treatment is considered small and does not warrant specific screening to any extent greater than age-appropriate general population health screening.

(Weak recommendation, Low-quality evidence)

Most long-term follow-up studies variably report a very low risk of secondary malignancies (bone and soft tissue malignancies, including breast, colorectal, kidney, and salivary cancers, and leukemia) in long-term survivors (762,763). A meta-analysis of two large multicenter studies showed that the RR of second malignancies was significantly increased at 1.19 (95% CI 1.04–1.36, \( p < 0.010 \)), relative to thyroid cancer survivors not treated with RAI, although the absolute increase in second primary malignancy risk attributable to RAI treatment is considered small and does not warrant specific screening to any extent greater than age-appropriate general population health screening.

(Weak recommendation, Low-quality evidence)

Published data indicate that when administered activities are selected to remain below 200 cGy to the bone marrow, minimal transient effects are noted in white blood cell and platelet counts (955). However, persistent mild decrements in white blood cell count and/or platelets are seen in some patients who have received multiple RAI therapies. Further, radiation to the bone marrow is impacted by several factors, including renal function. The kidneys are a major means of iodine excretion from the body, and physiologic radioisotope study research in nonthyroidectomized individuals has shown that renal impairment significantly reduces RAI excretion (994).

[C35] How should patients be counseled about RAI therapy and pregnancy, breastfeeding, and gonadal function?

- **RECOMMENDATION 88**

Women of childbearing age receiving RAI therapy should have a negative screening evaluation for pregnancy prior to RAI administration and avoid pregnancy for 6–12 months after receiving RAI.

(Strong recommendation, Low-quality evidence)

Radioactive iodine should not be given to nursing women. Depending on the clinical situation, RAI therapy could be deferred until lactating women have stopped breastfeeding or pumping for at least 3 months. A diagnostic \( ^{123} \text{I} \) or low-dose \( ^{131} \text{I} \) scan should be considered in recently lactating women to detect breast uptake that may warrant deferral of therapy.
(Strong recommendation, Moderate-quality evidence)

- **RECOMMENDATION 90**

  Men receiving cumulative RAI activities ≥400 mCi should be counseled on potential risks of infertility.

  (Weak recommendation, Low-quality evidence)

  Women about to receive RAI therapy should first undergo pregnancy testing. Gonadal tissue is exposed to radiation from RAI in the blood, urine, and feces. Temporary amenorrhea/oligomenorrhea lasting 4–10 months occurs in 20%–27% of menstruating women after ¹³¹I therapy for thyroid cancer. Although the numbers of patients studied are small, long-term rates of infertility, miscarriage, and fetal malformation do not appear to be elevated in women after RAI therapy (995–997).

  One recent large retrospective cohort study showed that use of RAI was associated with delayed childbearing and decreased birthrate in later years, although it is unclear if this is due to reproductive choice or reproductive health (998). Another large retrospective study suggested that pregnancy should be postponed for 1 year after therapy because of an increase in miscarriage rate (999), although this was not confirmed in a subsequent study (1000).

  Ovarian damage from RAI therapy may result in menopause occurring approximately 1 year earlier than in the general population, but this result was not associated with cumulative dose administered or the age at which the therapy was given (1001). Radioiodine is also significantly concentrated in lactating breast tissue (1002). Therefore, RAI should not be given to women who are breastfeeding (1003). A diagnostic ¹²³I or low-dose ¹³¹I scan can be employed in recently lactating women to detect breast uptake that may warrant deferral of therapy (662). Dopaminergic agents might be useful in increasing breast exposure in recently lactating women; although caution should be exercised given the risk of serious, albeit rare, side effects associated with their routine use to suppress postpartum lactation (1003,1004).

  In men, RAI therapy may be associated with a temporary reduction in sperm counts and elevated serum follicle-stimulating hormone levels (1005,1006). Higher cumulative activities (500–800 mCi) in men are associated with an increased risk of persistent elevation of serum follicle-stimulating hormone levels, but fertility and risks of miscarriage or congenital abnormalities in subsequent pregnancies are not changed with moderate RAI activities (∼200 mCi) (1007,1008). Permanent male infertility is unlikely with a single ablative activity of RAI, but theoretically there could be cumulative damage with multiple treatments. It has been suggested that sperm banking be considered in men who may receive cumulative RAI activities ≥2400 mCi (1006). Gonadal radiation exposure is reduced with good hydration, frequent micturition to empty the bladder, and avoidance of constipation (1009). Some specialists recommend that men wait 3 months (or one full sperm cycle) to avoid the potential for transient chromosomal abnormalities.

  **[C36] How is RAI-refractory DTC classified?**

- **RECOMMENDATION 91**

  Radioiodine-refractory structurally evident DTC is classified in patients with appropriate TSH stimulation and iodine preparation in four basic ways: (i) the malignant/metastatic tissue does not ever concentrate ¹³¹I outside the thyroid bed at the first therapeutic WBS, (ii) the tumor tissue loses the ability to concentrate RAI after previous evidence of RAI-avid disease (in the absence of stable iodine contamination), (iii) RAI is concentrated in some lesions but not in others; and (iv) metastatic disease progresses despite significant concentration of RAI. When a patient with DTC is classified as refractory to RAI, there is no indication for further RAI treatment.

  (Weak recommendation, Low-quality evidence)

  The prognosis of patients with DTC is usually favorable, even when metastatic RAI-avid disease is present. For this reason, ¹³¹I is considered the gold standard in the treatment of metastatic disease. However, many DTC patients with advanced disease do not respond or become refractory to ¹³¹I, with some of these patients dying within 3–5 years, but there are also long-term survivors with very slowly progressive disease.

  Iodine ¹³¹I refractory DTC includes four categories of patients (1010). In the first category, the malignant/metastatic tissue does not concentrate ¹³¹I outside the thyroid bed at the first therapeutic WBS. There is no evidence in these patients that further treatment with RAI may be of any benefit. Furthermore, patients with measurable disease with an absence of ¹³¹I uptake on subsequent diagnostic WBS may also be considered refractory because even when uptake is seen on posttherapy scan, it will likely have limited benefit. In the second category, the tumor tissue loses the ability to concentrate ¹³¹I (in the absence of stable iodine contamination) after previous evidence of uptake. This often occurs in patients with large and multiple metastases and is due to the eradication by ¹³¹I treatment of differentiated cells able to concentrate RAI but not of poorly differentiated cells that do not concentrate ¹³¹I and that generally will progress. In the third category, ¹³¹I uptake is retained in some lesions but not in others; this pattern is frequent in patients with multiple large metastases as shown by ¹²³I studies on PET scan (918) and by comparing results of ¹⁸FDG-PET scan with ¹³¹I WBS. In these patients, progression is likely to occur in metastases without uptake (in particular when ¹⁸FDG uptake is present), and RAI treatment will not be beneficial on the overall outcome of the disease. Some patients may have predominantly RAI-avid disease with only a few lesions that do not concentrate RAI. While these patients technically meet the definition of RAI-refractory disease, they may benefit from a combination of RAI therapy for the majority of the lesions and directed therapy (section [C38]; Recommendation 93) for those few lesions that do not concentrate RAI. In the fourth category, metastatic disease progresses despite significant uptake of ¹³¹I (new lesions, progressive growth of lesions, continual rise in serum Tg within months of RAI therapy). The definition of any of these possibilities depends on imaging modalities, including a posttherapy ¹³¹I WBS combined with other imaging modalities, such as CT scan, MRI, or ¹⁸FDG PET/CT. The criteria for RAI-refractory disease remain somewhat controversial. Future studies will hopefully help to refine this definition.

  Predictive factors for tumor response to RAI treatment are indeed the presence of ¹³¹I uptake by tumor, and among those patients with RAI uptake, younger age, well-differentiated histotype, small metastases, and low ¹⁸FDG uptake. Indeed these
parameters are closely interrelated (823,824,844) and may permit predicting the outcome of RAI treatment at the time of the discovery of the metastases. About two-thirds of patients with metastases demonstrate $^{131}$I uptake in their metastases, and only half of them will be cured with repeated courses of $^{131}$I treatment.

Less clear is the case of patients with visible RAI uptake in all the known lesions, who are not cured despite several treatment courses but whose disease remains stable and does not progress according to RECIST criteria. It is controversial whether these patients should be considered $^{131}$I refractory and whether RAI treatment should be abandoned in favor of other treatment modalities. The probability to obtain a cure with further treatments is low and side effects may increase, including the risk of secondary cancers and leukemias (762,763,844,961). Several parameters should then be taken into consideration for the decision to continue treatment with RAI, including response to previous treatment courses, high or significant $^{131}$I uptake after previous treatment courses, low $^{18}$FDG uptake in tumor foci, and limited side effects from the RAI therapy.

[C37] **Which patients with metastatic thyroid cancer can be followed without additional therapy?**

- **RECOMMENDATION 92**

  (A) Patients with $^{131}$I-refractory metastatic DTC that is asymptomatic, stable, or minimally progressive who are not likely to develop rapidly progressive, clinically significant complications and do not have indications for directed therapy can be monitored on TSH-suppressive thyroid hormone therapy with serial radiographic imaging every 3–12 months.

  (Weak recommendation, Low-quality evidence)

  (B) $BRAF$ or other mutational testing is not routinely recommended for prognostic purposes in patients with RAI-refractory, progressive, locally advanced, or metastatic DTC.

  (Weak recommendation, Moderate-quality evidence)

Patients with $^{131}$I-refractory metastatic DTC often have an indolent clinical course, with no apparent symptoms or adverse impact from their disease burden for many years. In the absence of tolerable therapies with significant likelihood of inducing durable complete remission or improving overall survival, treatment should be limited to interventions that prevent morbidity or palliate symptoms. Thus, once $^{131}$I refractory metastatic disease is identified, attention should be directed toward (i) determining the extent of metastatic disease by staging imaging studies such as CT, $^{18}$FDG PET/CT, or MRI as described in sections [C9–C13]; (ii) assessing degree of current or potential symptoms from the disease; (iii) understanding the comorbidities that might influence the choice of therapies for the metastatic disease; and (iv) determining the rate of progression of radiographically evident lesions. Serial assessment of the size and development of metastatic lesions can be enhanced by applying criteria similar to RECIST, as commonly used to assess tumor response in clinical trials (1011). Representative soft tissue metastatic lesions, typically >1 cm, are identified as “targets” on cross-sectional imaging, with the longest diameter of each lesion measured. Disease extent can be considered to be stable or minimally progressive if the sum of the longest diameters of the target lesions increases <20% in the absence of new metastatic foci on sequential imaging over 12–15 months of follow-up. No study has identified an appropriate frequency for repeat imaging, but it is reasonable to repeat the staging imaging studies within 3–12 months based on the disease burden and locations of lesions (953,1012). More frequent assessment could be considered once metastatic disease is identified and/or in response to intercurrent patient symptoms, and less frequent imaging performed once a pattern of stability is identified. The development of progressive disease, either by RECIST assessment identifying at least 20% increase in sum of longest diameters of target lesions, the appearance of significant new metastatic lesions, or development of disease and requiring more aggressive therapy solely on the basis of rising levels of $Tg$; accelerating increases in $Tg$ levels should, however, lead to consideration of more frequent and comprehensive imaging in efforts to identify previously occult structural correlates.

In each of two recent international, randomized phase III trials of kinase inhibitors for therapy of progressive, RAI-refractory, metastatic DTC, patients whose tumors contained the $BRAF^{V600E}$ mutation who were randomized to the placebo treatment arms experienced similar progression-free survival compared with those whose tumors did not harbor a $BRAF$ mutation (1013,1014). Thus, $BRAF$ mutation status does not appear to be of prognostic value once patients reach this degree of advanced disease that would have qualified them for these clinical trials, that is, RAI-refractory, locally advanced, or metastatic DTC that has progressed by RECIST within 13 or 14 months.

[C38] **What is the role for directed therapy in advanced thyroid cancer?**

- **RECOMMENDATION 93**

  (A) Both stereotactic radiation and thermal ablation (RFA and cryoablation) show a high efficacy in treating individual distant metastases with relatively few side effects and may be considered as valid alternatives to surgery.

  (Weak recommendation, Moderate-quality evidence)

  (B) Stereotactic radiation or thermal ablation should be considered prior to initiation of systemic treatment when the individual distant metastases are asymptomatic or at high risk of local complications.

  (Strong recommendation, Moderate-quality evidence)

Several local treatment modalities other than surgery may be used to treat brain, lung, liver, and bone lesions from thyroid carcinoma. These local treatment modalities should be considered prior to initiation of systemic treatment when the individual distant metastases are asymptomatic or at high risk of local complications. They may also be helpful in case
of progression in a single lesion in patients with otherwise controlled disease during systemic treatment. In these patients, benefits can be achieved in preventing local complications, in improving symptoms such as pain, in delaying the initiation of systemic treatments, and even in improving survival. These techniques can be a less aggressive alternative to surgery and may be indicated in cases of lung metastases associated with insufficient respiratory reserve, poor patient clinical status, or after multiple previous surgical resections, local recurrence at the site of previous surgery, or refusal of additional surgery.

In selected patients, the techniques may be an alternative to surgery as first-line treatment, and they may induce local tumor control with a similar efficacy to surgical resection. Of interest, long-term benefits in terms of disease control have been reported in patients with a single or few metastases and in whom the disease is slowly progressive.

Interventional radiology (thermal ablation and cement injections), SBRT, or intensity modulated radiotherapy are the most frequently used techniques. The main principle of these techniques is to selectively treat the lesion, be minimally invasive, and be well tolerated with relatively few side effects. The indications and the feasibility of each technique depend on the location and the size of the lesion to be treated. Experience with metastases from thyroid cancer is scarce, and most available data have been obtained in patients with metastases from nonthyroid cancers.

SBRT allows for delivering high radiation doses in few fractions to the target tumoral lesion with a high degree of precision, minimizing the radiation of normal surrounding tissue. It has been used in several trials to treat brain, liver, lung, and bone metastases.

For patients with few (one to three) brain metastases, SBRT is as effective as surgery and can be repeated in case of appearance of new brain lesions. It is usually well tolerated, and brain necrosis that occurred in less than 10% of cases is usually limited and had no clinical consequences. The patient outcome depends mostly on the progression rate of extracerebral lesions (1015).

Data on lung and liver metastases are available only in retrospective studies on low numbers of patients and with a median follow-up of less than 1 year in most cases and in one prospective study (1016). This study included patients with many different primary tumors, including 10% of the patients having thyroid cancer. They showed a local control rate ranging from 63% to 98% in lung lesions and from 57% to 100% in liver lesions, with a cumulative dose delivered ranging from 20 to 75 Gy in 5–15 fractions. The local tumor control seems to be long lasting with complete response ranging from 70% to 90% at 2–3 years. Furthermore, rare (<5%) grade 3–4 toxicities (pneumonitis, pleural effusion, intestinal complications) were reported (1017). These toxicities are much less common that those associated with percutaneous treatment modalities. The local control rate seems to be dose-related, but the optimal protocol for SBRT is still not established.

Concerning bone lesions, radiotherapy plays an important role because it can complement surgery in case of incomplete resection or be used alone for pain relief or palliation. In many cases, usual radiotherapy can be given. However, the major limitation of radiotherapy in spine lesions is the cumulative dose to the spinal cord. SBRT compared to standard radiotherapy demonstrated a higher efficacy on tumor control and for limiting radiation to the spinal cord, especially in patients who need to be re-irradiated. A local tumor control rate of bone lesions for SBRT ranging from 88% to 100% was reported especially for lesions that were previously surgically resected, with a pain relief rate of 30%–83%. SBRT protocols differed among studies, with a maximum of 30 Gy administered in one to five fractions but a single dose of 12.5–15 Gy seems to achieve similar results (1018). Spinal myelopathy or vertebral fractures are the most important side effects, especially in case of large-volume lesions.

Percutaneous thermal ablation is aimed at destroying tumor foci by increasing (RFA) or decreasing (cryoablation) temperatures sufficiently to induce irreversible cellular damages. RFA is performed for liver, lung, and bone tumor foci. Clinical trials showed high efficacy of RFA for liver lesions, especially from colorectal cancer with long-term local disease control ranging from 40% to 80%, depending on lesion size, and a prolonged overall survival in treated patients (1019,1020).

In a total of 100 lung lesions including primary lung tumors and metastases, RFA was both efficient and well tolerated, with a complete tumor control rate of 93% at 18 months (1021). A multicenter prospective trial on 183 lung metastases from cancer other than colorectal showed a complete response rate of 88% at 1 year and an overall survival of 92% and 64% at 1 year and at 2 years, respectively (1022). Recurrence occurs more frequently in lesions >3 cm and can be detected early after ablation (within 3 months) by 18FDG PET/CT (824,1016). Cases of delayed recurrence have also been reported, and long-term follow-up is needed. RFA may be considered for lung lesions <3 cm of diameter, without soft tissue or mediastinum invasion and without contact with large vessels. Furthermore, repeated treatments can be performed on the same lesion and multiple lesions can be treated in the same patient.

RFA or cryoablation of bone lesions showed promising results with rapid (1–7 days) and long-lasting pain control (1023,1024). Cryoablation is a safe technique to treat or to stabilize bone lesions, is frequently associated with cementoplasty to consolidate the bone and avoid subsequent complications, and can treat larger lesions than RFA.

Local disease control was achieved in the few reported cases of lung and bone metastases from thyroid cancer treated by thermal ablation (883,1016,1025). The association of cryoablation and cementoplasty seems promising in purely lytic bone metastases from thyroid cancer. Furthermore, in three patients with liver metastases from thyroid cancer (two MTC and one FTC) RFA was feasible and reduced local symptoms (1026). Multidisciplinary treatment for metastases from thyroid cancer especially for bone metastases—including, for example, thermal ablation (RFA or cryoablation) and/or cementoplasty associated with systemic treatment (RAI, bone-directed agents, or chemotherapy)—can improve the patient’s quality of life by reducing pain and prolonging time to skeletal events, delaying initiation of systemic treatment, and even improving patient survival (1027).

The toxicity of thermal ablation is generally low, but pneumothorax or pleural effusion was observed in up to 50% of RFA procedures for lung lesions, but it rarely required further treatments. Local pain or transient neurological deficit or vertebral fracture can occur in case of ablation of bone.
lesions (5%-6%), and intestinal perforation, abdominal pain, or intraperitoneal bleeding was observed following ablation of liver lesions (1021,1028).

Published experience using thermal ablation and stereotactic radiation in thyroid cancer patients is limited, and recommendations are currently based on more robust evidence in other solid tumors. Randomized prospective studies comparing the efficacy and tolerability of these different techniques are lacking, and their choice in clinical practice is based on local experience, lesion location as well as patient status and preference.

[C39] Treatment of brain metastases

- **RECOMMENDATION 94**

While surgical resection and stereotactic EBRT are the mainstays of therapy for CNS metastases, RAI can be considered if CNS metastases concentrate RAI. If RAI is being considered, stereotactic EBRT and concomitant glucocorticoid therapy are recommended prior to RAI therapy to minimize the effects of a potential TSH-induced increase in tumor size and RAI-induced inflammatory response.

*(Weak recommendation, Low-quality evidence)*

Brain metastases typically occur in older patients with more advanced disease and are associated with a poor prognosis (933). Surgical resection and stereotactic EBRT are the mainstays of therapy (933,1029,1030). There are few data showing efficacy of RAI. Stereotactic radiation therapy is preferred to whole-brain radiation because life expectancy in patients with brain metastases may be prolonged, and stereotactic radiation induces less short- and long-term toxicity compared with whole-brain radiation (fatigue, headache, cognitive decline, and behavioral changes), and it may be effective even in patients with multiple brain lesions.

[C40] Who should be considered for clinical trials?

- **RECOMMENDATION 95**

Patients should be considered for referral to participate in prospective therapeutic clinical trials based upon specific eligibility requirements for given studies and the likelihood that the patient may or may not benefit from study participation. Clinicians considering referral of patients for trials should review available treatment options and eligibility criteria, preferably through discussions with trial center personnel and review of trial materials at the website www.clinicaltrials.org.

*(Strong recommendation, Moderate-quality evidence)*

A therapeutic clinical trial is a systematic investigation of the effectiveness and safety of a potential new or modified treatment or combination of treatments, potentially including medications, surgery, radiation therapy, and/or other novel or revised approaches. A broad variety of such trials may exist at any given time, which can generally be identified through online databases such as www.clinicaltrials.org, best supplemented with direct contact with the institutions conducting trials of particular interest so as to ensure trial availability and patient eligibility. There is limited evidence that enrollment into clinical trials is associated with lower overall cancer-specific mortality for patients with common cancers, even within contexts in which approved and “standard of care” therapies already exist (1031). The reasons for this association are unclear, but there is no evidence to suggest that trial participation is deleterious to patient outcomes, and it may be beneficial. Participation in a clinical trial should be considered in any situation wherein there exists no effective or proven standard of care, or when a standard of care is being compared with a promising new or investigational therapy. Adjunct therapy trials may be appropriate for patients at high risk for disease recurrence following primary treatment who wish to pursue aggressive therapy. For patients with RAI-refractory carcinoma that is locally advanced or metastatic, clinical trials may be appropriate in the setting of disease that is considered progressive by RECIST criteria, especially if progression occurred after use of an approved kinase inhibitor such as lenvatinib or sorafenib, and/or if approved therapies are otherwise unaffordable to a specific patient. Patients demonstrated to have specific therapeutically targetable tumor alterations, such as BRAF mutations or PAX8/PPARγ or ALK rearrangements, can also be considered for trials testing therapies specifically targeting these alterations. However, given the indolent nature of metastatic disease in most patients, therapeutic clinical trial participation should not be considered for patients with stable, asymptomatic metastatic disease unless agents with significant likelihood of complete remission, prolongation of survival, or biologic impact such as redifferentiation that could sensitize to definitive therapy are available (see section [C37]).

[C41] What is the role of systemic therapy (kinase inhibitors, other selective therapies, conventional chemotherapy, bisphosphonates, denosumab) in treating metastatic DTC?

Systemic therapeutics of several types in selected clinical contexts appear to provide clinical benefit in treating metastatic DTC (1032). Benefit has been demonstrated in the form of improved progression-free survival (delay in time to disease progression or death) in three randomized, double-blinded, placebo-controlled clinical trials: vandetanib (1033), sorafenib (1013), and lenvatinib (1034). Benefit has also been demonstrated in the form of induced durable tumor regression (1035–1037).

However, randomized clinical trial data are not yet available to address many additional critical questions, including effects of systemic therapies of various types on survival and quality of life, or to address critical issues of optimal patient selection/inclusion/exclusion criteria for therapy and duration of treatment.

To date, no clinical trial has demonstrated an overall survival advantage or improved quality of life from use of any therapy in RAI-refractory DTC (1013,1034). Consequently, therapeutic decisions are presently based upon the convergence of expert opinion and patient preference/philosophy, thus emphasizing the critical need to address the above questions definitively through clinical trials. It is therefore important to involve highly skilled clinicians familiar with RAI-refractory disease and these systemic therapies in decision-making until definitive guidelines can be developed based upon more rigorous data. As a guide, evidence-based recommendations with expert consensus have been recently published (953).

It should also be highlighted that, broadly construed, systemic therapy encompasses not only more recently emerging
“targeted” approaches, but also historical “mainstay” therapies including TSH suppression and RAI. Although more “novel” approaches have attracted attention recently, it is important to optimally apply fundamental approaches. In this regard, therapeutic RAI should also be used to optimal effect prior to the initiation of more recent/novel therapies.

To accomplish this requires attention to detail, including ensuring adequate TSH stimulation, patient adherence to low-iodine pre-RAI therapy dietary restrictions, and avoidance of proximal preceding iodine contamination from IV contrast agents, with verification by urinary iodine concentration measurements in selected cases. In this context, occasional patients previously declared “RAI-refractory” can instead be found to have RAI-responsive disease when such details are attended to if previously neglected.

Also important is the consideration of alternatives to the use of systemic therapy, such as the application of surgery or other localized approaches (including radiation therapy or thermal ablation approaches). In patients in whom the threats imposed by cancer are more localized, directed approaches may have greater potential to control localized disease and symptoms compared to systemic therapies, especially given the absence of data to indicate survival benefit or improved quality of life from the application of systemic therapies in advanced RAI-refractory DTC.

It is also critically important to ensure that the disease prompting therapy represents metastatic thyroid cancer. In particular, because pulmonary nodules attributable to benign causes are common, the presence of pulmonary nodules does not in and of itself justify the application of systemic therapy. Thus, in cases of diagnostic uncertainty in which the result would have definitive therapeutic implications, biopsy is required, especially when Tg levels are low/unhelpful (such as in the presence of anti-Tg antibodies). Conversely, stable, asymptomatic pulmonary nodules of a few millimeters in size likely do not justify invasive assessment or systemic therapy.

The introduction of systemic therapy requires that both the clinician and the patient agree that clinical benefits are expected to exceed risks for that individual patient. The problem in this determination, however, is that it is often very difficult to precisely define such risks and benefits because they vary greatly depending upon patient context and they are often poorly articulated in the literature. It is also critical to weigh not just risks of death and injury, but also risks of systemic therapies on quality of life, especially as viewed by a particular patient considering treatment. Hence, the decision is not based solely on benefits and risks of therapy, but also on patient value judgments. Issues of risks and benefits are reviewed in this context in conjunction with each therapeutic modality below. Finally, it is important that the involved care team (physicians, physician assistants, nurse practitioners, nurses) be experienced in the use and management of toxicities associated with these therapies.

Kinase inhibitors that are FDA approved for differentiated thyroid carcinoma or other available kinase inhibitors (preferably within the context of therapeutic clinical trials) can be considered since the impact of these agents on overall survival and quality of life remains to be defined.

**Weak recommendation, Moderate-quality evidence**

(B) Patients who are candidates for kinase inhibitor therapy should be thoroughly counseled on the potential risks and benefits of this therapy as well as alternative therapeutic approaches including best supportive care. Appropriate informed consent should be obtained and documented in the medical record prior to initiation of any therapy, regardless of whether the patient is being treated in the context of a clinical trial.

**Strong recommendation, Low-quality evidence**

Cytotoxic chemotherapy has historically produced disappointing results in patients with DTC (1038). Kinase inhibitors, many of which share the common target of the VEGF receptor (VEGFR) (e.g., sorafenib, pazopanib, sunitinib, lenvatinib, axitinib, cabozantinib, and vandetanib), have recently emerged as highly promising therapies for metastatic RAI-refractory DTC (1032). Kinase inhibitors, however, are associated with numerous adverse effects including diarrhea, fatigue, induced hypertension (requiring initiation of antihypertensive therapy in about half of all previously normotensive individuals), hepatotoxicity, skin changes, nausea, increased LT4 dosage requirement, changes in taste, and weight loss. These potential side effects have high probability of negatively impacting quality of life and/or necessitating dosage reductions in many patients and treatment discontinuation in up to 20% of patients. Furthermore, these agents are also associated with more serious and potentially fatal risks including from thrombosis, bleeding, heart failure, hepatotoxicity, gastrointestinal tract fistula formation, and intestinal perforation (1039). Overall, the risk of therapy-related death in cancer patients treated with oral kinase inhibitors is about 1.5%–2% (RR = 2.23, p < 0.023, compared with randomized placebo treated control cancer patients) based upon meta-analysis of results from 10 recently published randomized trials conducted in several cancers (1039). In the setting of thyroid cancer, lenvatinib (now approved in the United States for use in RAI-refractory DTC) was associated with severe toxicities in 75% of patients and therapy-attributed mortality in 2.3% of patients (1034). While risk of drug-related death is relatively low, the knowledge of potentially fatal therapeutic outcomes should prompt considerable restraint in the use of kinase inhibitors, especially in patients who are symptomatic and/or with stable or slowly progressive disease and in patients who can otherwise be effectively treated using directed therapies.

Three randomized placebo-controlled clinical trials (phase 2, vandetanib; phase 3, sorafenib and lenvatinib) had been published by the time of the writing of these guidelines, each demonstrating delayed time to disease progression among kinase inhibitor–treated patients relative to those treated with a placebo (1013,1033,1034). On this basis, sorafenib and lenvatinib were approved for use in the United States and the European Union for patients with advanced RAI-refractory
Diffuse disease progression as opposed to focal progression associated with progression-free survival prolonged by 5 months, with <15% objective response rates, but with no improvement in overall survival demonstrated to date. Assessments comparing sorafenib to placebo outcomes demonstrated overall lower quality of life among sorafenib-treated patients relative to those treated with placebo, despite improved progression-free survival. Lenvatinib therapy was associated with longer prolonged median progression-free survival by 14.7 months compared with placebo, with a RECIST response rate of 65%, with some complete responses also reported. Despite these very encouraging results and regulatory approval, however, no statistically significant impact of lenvatinib (or any other kinase inhibitor) therapy on overall survival has yet been observed (1034). However, survival data in this context, where all three randomized and placebo-controlled DTC trials allowed crossover of placebo-treated patients to kinase inhibitor, must be interpreted with great caution. The crossover trial designs effectively “contaminated” overall survival data, with many placebo-treated patients later crossing over to receive open-label kinase inhibitor. Hence, the absence of a proven effect of kinase inhibitors on overall survival in these trials can best be interpreted as indicating that, among study populations, a delay in initiation of kinase inhibitor therapy until observation of RECIST disease progression among initially placebo-treated patients did not adversely affect overall survival.

Additional VEGFR-directed kinase inhibitors including axitinib, pazopanib, cabozantinib, and sunitinib also have activity in metastatic DTC based upon phase 2 therapeutic trials (1035–1037). No multi-arm comparison “superiority” phase 3 trial data in DTC are available to inform decision-making with regard to selection among available and mostly phase 3 trial data in DTC are available to inform decision-making with regard to selection among available and mostly similarly targeted kinase inhibitors. Recent regulatory approval in DTC and more abundant outcome data in response to lenvatinib and sorafenib therapy in DTC (1013,1034,1040–1043) prompt prominent consideration of either lenvatinib or sorafenib as a first-line therapy.

Despite encouraging trial results, the extent to which any kinase inhibitor may prolong overall survival and in which contexts remains undefined. In some situations, however, the decision to initiate kinase inhibitor therapy is straightforward. For example, patients with oxygen dependence attributable to DTC lung metastases are not only adversely affected by their cancers, but also imminently threatened. Such patients have potential to gain symptomatic benefit even in the absence of definite extension of life. At the other extreme, many patients with metastatic DTC are asymptomatic from their cancers and are not expected to be threatened by their cancers in the foreseeable future. These patients should remain on TSH-suppressive treatment as their primary therapeutic intervention rather than be exposed to kinase inhibitor therapy. For patients with disease between these two extremes, it is critical that risks and benefits of therapy, anticipated side effects, goals, patient context, and the therapeutic philosophy of each particular patient be thoroughly vetted so as to best individualize therapy. This decision is ideally made within the context of specialized centers with comprehensive knowledge of the natural history of the disease and of the effects of available therapies.

As a general “expert consensus” guideline, structurally progressive, symptomatic, and/or imminently threatening DTC (wherein disease progression is expected to require intervention and/or to produce morbidity or mortality in <6 months) that is RAI refractory and not amenable to satisfactory control using directed approaches (e.g., surgery, radiation therapy, thermal ablation) should prompt consideration of kinase inhibitor therapy. Specific characteristics impacting decision-making for starting such therapy are outlined in Table 16.

Patients who are candidates for kinase inhibitor therapy should be thoroughly counseled with regard not only to

<table>
<thead>
<tr>
<th>Factors favoring kinase inhibitor therapy</th>
<th>Factors discouraging kinase inhibitor therapy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Imminently threatening disease progression expected to require intervention and/or to produce morbidity or mortality in &lt;6 months (e.g., pulmonary lesions or lymphadenopathy likely to rapidly invade airways, produce dyspnea, or cause bronchial obstruction).</td>
<td>Comorbidity including</td>
</tr>
<tr>
<td>Symptomatic disease (e.g., exertional dyspnea, painful unresectable adenopathy), not adequately addressable using directed therapy.</td>
<td>• Active or recent intestinal disease (e.g., diverticulitis, inflammatory bowel disease, recent bowel resection)</td>
</tr>
<tr>
<td>Diffuse disease progression as opposed to focal progression (e.g., in multiple lung metastases, as opposed to a few growing lesions)</td>
<td>• Liver disease</td>
</tr>
<tr>
<td></td>
<td>• Recent bleeding (e.g., ulcer/GI bleed) or coagulopathy</td>
</tr>
<tr>
<td></td>
<td>• Recent cardiovascular event(s) (e.g., CVA, MI)</td>
</tr>
<tr>
<td></td>
<td>• Recent tracheal radiation therapy (this is associated with increased risks of aerodigestive fistula with kinase inhibitor therapy)</td>
</tr>
<tr>
<td></td>
<td>• Cachexia/low weight/poor nutrition</td>
</tr>
<tr>
<td></td>
<td>• Poorly controlled hypertension</td>
</tr>
<tr>
<td></td>
<td>• Prolonged QTc interval/history of significant arrhythmia (includes ventricular and bradycardias)</td>
</tr>
<tr>
<td></td>
<td>• Untreated brain metastases (controversial)</td>
</tr>
<tr>
<td></td>
<td>• Recent suicidal ideation (suicide has been reported in depressed patients receiving TKIs)</td>
</tr>
<tr>
<td></td>
<td>Life expectancy based upon other comorbidities estimated to be too brief to justify systemic therapy</td>
</tr>
</tbody>
</table>

*Bone metastases are often poorly responsive to kinase inhibitor therapy (see Bone-Directed Agents in section [C47]). GI, gastrointestinal; CVA, cerebrovascular accident; MI, myocardial infarction; TKI, tyrosine kinase inhibitor.
potential benefits, but also about potential side effects and risks of therapy and alternative therapeutic approaches including best supportive care, as should be the case with any medical therapeutic decision-making (1044). Such extensive and comprehensive discussions are particularly important to undertake in the context of kinase inhibitor therapy because of the high probability of side effects of these agents and because of their presently uncertain effects on patient overall survival and quality of life (1045).

Another critical and complex question often faced in treating patients with kinase inhibitor therapy centers on when treatment should be discontinued once initiated. In general, therapy should be continued so long as net benefit exceeds net detriment. In case of slow RECIST disease progression after significant tumor response, treatment may be maintained so long as overall disease control is maintained providing that toxicities are manageable. Progression rate should be taken into account, and when global progression is rapid, therapy should be discontinued. However, there are times when focal/oligometastatic disease progression amenable to directed therapies is noted. In such instances, superimposed loco-regional therapies in the setting of maintained systemic therapy can sometimes be in the best interest of a particular patient. For instance, in the setting of regressed lung metastases and yet progression at a solitary bone site, maintained systemic therapy with superimposed directed radiation therapy may be reasonable.

[C43] Patients for whom first-line kinase inhibitor therapy fails

- **RECOMMENDATION 97**
  Patients who have disease progression while on initial kinase inhibitor therapy without prohibitive adverse effects should be considered for second-line kinase inhibitor therapy. Ideally, such therapy should be undertaken within the context of therapeutic clinical trials.
  
  *(Weak recommendation, Low-quality evidence)*

- **RECOMMENDATION 98**
  DTC patients who progress through first-line kinase inhibitor therapy commonly respond to a second similarly targeted agent, and thus they should be considered candidates for second-line kinase inhibitor therapy (1014,1046). Hence, the selection of an agent for initial therapy may be less critical in some senses because many patients will go on to receive several kinase inhibitors over their disease courses. In the near future, other treatment modalities, such as RAI resensitization therapy, immunotherapy, or drugs directed to other targets, may also offer additional therapeutic options.
  
  *(Strong recommendation, Low-quality evidence)*

- **RECOMMENDATION 99**
  Agents without established efficacy in DTC should be used primarily within the context of therapeutic clinical trials.
  
  *(Strong recommendation, Low-quality evidence)*

A variety of novel agents have been and/or are being tested as candidate therapeutics in progressive metastatic RAI-refractory DTC (1032,1048). At the time of the writing of these guidelines, however, only kinase inhibitors have shown sufficient promise to consider use other than within the context of therapeutic clinical trials. Agents of particular interest for further testing include BRAF kinase inhibitors because PTC frequently harbors the constitutively activating \( BRAF^{V600E} \) mutation and the inhibitors have already shown efficacy and been approved for use in \( BRAF \) mutant melanoma (1049–1051). Inhibitors of MEK kinase and other signaling pathways are also of considerable investigational interest. In addition, promising initial results in response to use of kinase inhibition to “resensitize” RAI-refractory tumors to RAI have been reported (1052). Data developed to date, however, do not yet favor the use of these novel approaches over use of VEGFR-directed kinase inhibitors unless within the context of therapeutic clinical trials or alternatively when used as “salvage” therapies after disease progression has occurred despite prior VEGFR-directed kinase inhibitor therapy.

- **RECOMMENDATION 100**
  Cytotoxic chemotherapy can be considered in RAI-refractory DTC patients with metastatic, rapidly progressive,
symptomatic, and/or imminently threatening disease not otherwise amenable to control through other approaches, including kinase inhibitors. Too few data exist to recommend specific cytotoxic regimens, and use within the context of a therapeutic clinical trial is preferred.

*(Weak recommendation, Low-quality evidence)*

Although doxorubicin was approved for use in thyroid cancer by the US FDA in 1974 and has some utility in anaplastic thyroid cancer, cytotoxic chemotherapy has historically produced disappointing results when used to treat RAI-refractory DTC (1038). Cytotoxic chemotherapy, however, may have selective benefit in patients unresponsive to kinase inhibitors and perhaps also in some patients with poorly differentiated thyroid cancer (1053). Data are limited and primarily anecdotal.

*(Strong recommendation, Moderate-quality evidence)*

Metastatic bone disease represents a particularly challenging clinical problem in patients with RAI-refractory DTC, especially given the high rate of multiple skeletal-related events in patients following detection of an initial bone lesion (1054). Patients with a small number of threatening and/or symptomatic bone lesions are generally best treated with focal approaches such as radiation therapy and/or surgery and/or thermoaolbation. Many patients, however, suffer from diffuse progression of bone metastases that are not amenable to effective control using focal therapies alone. In such patients, focal therapy to symptomatic lesions or lesions at high risk of complications may be beneficial and should be performed before initiation of systemic treatment.

Unfortunately, kinase inhibitors appear to be less effective in controlling bone metastatic disease in comparison to disease at other soft tissue sites such as lungs and lymph nodes. Progression of bone metastases while on kinase inhibitor therapy commonly occurs despite maintained benefit with respect to disease at other metastatic sites. Hence, kinase inhibitor therapy cannot be relied upon to control diffuse bone metastases in many patients with RAI-refractory DTC.

In other solid tumors, bone-directed therapeutics including bisphosphonates (especially zoledronic acid) and the RANK ligand-directed agent denosumab have been shown to delay time to occurrence of subsequent skeletal-related adverse events (fracture, pain, neurologic complications) and to improve symptoms, and these agents may provide benefits for patients

---

**Table 17. Potential Toxicities and Recommended Screening or Monitoring Approaches in Patients Started on Kinase Inhibitor Therapy**

<table>
<thead>
<tr>
<th>Toxicity</th>
<th>Recommended screening/monitoring</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hypertension</td>
<td>Frequent blood pressure monitoring, most critical during the first 8 weeks of therapy; if hypertension is induced, therapy should be individualized to patient response</td>
</tr>
<tr>
<td></td>
<td>- Note: effective and expeditious management of hypertension is critical - and may reduce potential for cardiotoxicity. If antihypertensive therapy is needed, calcium channel blockers (e.g., amlodipine) may be most effective.</td>
</tr>
<tr>
<td>Cutaneous/mucocutaneous toxicities</td>
<td>Careful patient reporting of rash/mouth sores, patient awareness and education related to increased potential for photosensitization/sunburn.</td>
</tr>
<tr>
<td>Hepatotoxicity</td>
<td>Serial assessment of alanine serum transferase (AST), alkaline phosphatase and bilirubin - most critical during the first 8 weeks of therapy</td>
</tr>
<tr>
<td></td>
<td>- Note: dose reduction of kinase inhibitor therapy is commonly required due to hepatic toxicity</td>
</tr>
<tr>
<td>Cardiotoxicity</td>
<td>ECG pretherapy and frequently during therapy</td>
</tr>
<tr>
<td></td>
<td>- Hold (or do not initiate) kinase inhibitor therapy if QTc &gt;480 ms</td>
</tr>
<tr>
<td></td>
<td>- Echocardiogram: elective, but recommended in any patient with cardiac history and especially important in patient with hypertension, symptoms consistent with congestive heart failure or coronary artery disease and in patients receiving sunitinib</td>
</tr>
<tr>
<td>Hypothyroidism</td>
<td>TSH should be assessed frequently, with levothyroxine dosage altered in response to rising TSH if observed</td>
</tr>
<tr>
<td>Nephrotoxicity</td>
<td>Serial serum creatinine, urine analysis with protein determination,</td>
</tr>
<tr>
<td>Hematological toxicities</td>
<td>Serial CBC/diff</td>
</tr>
<tr>
<td>Pancreatectis</td>
<td>Serial amylase</td>
</tr>
<tr>
<td>Teratogenicity</td>
<td>Pretherapy pregnancy testing and effective contraception in women and men of childbearing potential</td>
</tr>
</tbody>
</table>

CBC, complete blood count; ECG, electrocardiography.
with diffuse bone metastases (1055,1056). The determination of benefits across several tumor types suggests that they may be broadly generalizable, prompting FDA approval for their general use in patients with solid tumor bone metastases. Two small studies have suggested benefit from bisphosphonates specifically within the context of DTC bone metastases (951,1057).

Risks of bisphosphonates and RANK ligand–directed agents include hypocalcemia, which can be severe, prompting the recommended concomitant use of supplemental calcium and vitamin D therapy. Moreover, these agents also moderately increase the risk of nonhealing oral lesions and jaw osteonecrosis; thus, candidates for this therapy should undergo dental/oral surgical evaluation prior to their initiation so as to minimize these risks (1055).

Recent studies have overall shown equivalence or superiority of denosumab to zoledronic acid in delaying bone-related adverse events in several solid tumors, with similar risk of jaw osteonecrosis, greater incidence of hypocalcemia, and less nephrotoxicity (1058). However, no thyroid-specific data related to the efficacy of denosumab have yet been published. If the patient’s renal function is impaired, there is sometimes justification for beginning with RANK ligand–directed/denosumab therapy because this agent seems to produce fewer adverse renal effects, albeit the oral/jaw effects are similar to bisphosphonates in magnitude.

Expert consensus is that bone-directed therapy should be strongly considered in patients with multiple progressing and/or symptomatic bone metastases, likely best beginning with bisphosphonate/zoledronic acid (assuming calcium and renal function permit). Candidates for such therapy should be cleared by their dentist/oral surgeon prior to therapy initiation; in addition, calcium and vitamin D therapy should be ongoing in conjunction with any intended bone-directed therapeutic. In general, there is consensus that the administration of zoledronic acid therapy every 3 months (rather than every month) is a reasonable initial approach in terms of dosing interval, but randomized trial data are unavailable to definitively clarify this issue. Expert consensus is that bone-directed therapy should be used in the setting of diffuse bone metastases even if kinase inhibitor therapy is intended or ongoing.

[D1] DIRECTIONS FOR FUTURE RESEARCH

[D2] Optimizing molecular markers for diagnosis, prognosis, and therapeutic targets

Significant progress has been made over the last several years in understanding the genetic mechanisms of thyroid cancer and creating molecular tests for cancer diagnosis in thyroid nodules. This process is currently going through an accelerated phase, which is expected to continue into the future. The Cancer Genome Atlas (TCGA) and work from multiple research laboratories led to the identification of mutations and other driver genetic alterations in more than 90% of thyroid cancers, making it one of the best characterized human cancers from a genetic standpoint (1059). The TCGA focused only on PTC. Moreover, next-generation sequencing technologies may allow detection of most of these alterations in a limited cell sample obtained by FNA. Progress in identifying mutational, other genetic (gene expression, miRNA), and epigenetic markers of thyroid cancer is expected to result in a significantly improved accuracy of cancer detection in thyroid nodules as compared to the currently available clinical tests. If such progress continues, it is expected that future molecular tests will be able to predict the risk of cancer in thyroid nodules with high accuracy, dramatically reducing the uncertainty of indeterminate FNA cytology. Furthermore, as the cost of next-generation sequencing of human DNA continues to decrease and the analytical tools become more efficient, it should be expected that the cost of molecular testing will decrease, enabling cost-efficient utilization of testing.

Molecular markers are expected to have a significant impact on cancer prognostication. While the 
BRAF
status can be considered as a relatively sensitive prognostic marker for papillary cancer, it cannot be used in isolation for tumor prognostication. However, recent results obtained by broad tumor genotyping show that several specific molecular signatures (such as presence of several driver mutations, 
TP53
mutation, 
TERT
mutation in isolation or in combination with 
BRAF
) are found in a small fraction of well-differentiated papillary and follicular cancers and appear to be associated with more aggressive tumor behavior. It is expected that these molecular signatures will be confirmed and perhaps further improved in additional studies and will offer a more specific detection of well-differentiated thyroid cancers that have high risks of tumor recurrence and cancer-related mortality. Furthermore, research may identify how such data may inform therapeutic decision-making (e.g., surgical extent [if any], treatment with multi-kinase and specific kinase inhibitors or their combination). Discoveries of new gene mutations/rearrangements involved in the pathogenesis of thyroid cancer, such as those of 
ALK
and 
NTRK3
, are expected to offer new effective therapeutic targets. Finally, new therapeutic approaches to target genes commonly mutated in thyroid cancer, such as the 
RAS
genes, are in development and expected to enter clinical trials in the future. Prospective long-term outcome studies, ideally in the form of RCTs, will be needed to define the optimal surgical and postsurgical management of patients based on these molecular signatures. Such research may enable personalized, evidence-based care of patients with thyroid cancer across the disease trajectory.

[D3] Active surveillance of DTC primary tumors

Our Japanese colleagues have provided compelling data that an active surveillance management approach to papillary microcarcinoma is a safe and effective alternative to immediate surgical resection in properly selected patients (143,149).

Unfortunately, no clinical features or molecular abnormality in isolation can reliably differentiate the relatively small number of PTMC patients destined to develop clinically significant progression from the larger population of people that harbor indolent PTMCs that will not cause significant disease. Therefore, additional studies are needed to identify specific risk factors that would favor surgical resection over active surveillance.

Furthermore, additional studies are needed to define important management issues that arise during an active surveillance follow-up approach. These issues include the frequency of US evaluations required during follow-up, optimal TSH goals, the potential role of serum Tg in follow-up, and specific indications for surgical intervention (e.g., what measurements should be used to define a clinically significant
increase in the size of the primary tumor, what constitutes clinically significant lymph node metastases).

Finally, studies that examine decision-making and acceptability of an active surveillance approach to thyroid cancer in patients, family members, and clinicians are required to better understand how to implement this novel management approach outside of Japan. These studies should ideally be performed in the context of an Institutional Review Board–approved clinical trial.

[D4] Improved risk stratification

While the AJCC/UICC TNM staging provides valuable information with regard to disease-specific mortality, studies are needed to determine if inclusion of additional prognostic variables into the AJCC staging system could improve its predictive ability. Potential variables for consideration include the specific histology (well-differentiated thyroid cancer versus poorly differentiated thyroid cancer), molecular profile, size and location of distant metastases (pulmonary metastases versus bone metastases versus brain metastases), functional status of the metastases (RAI avid versus 18FDG-PET avid), and effectiveness of initial therapy (completeness of resection; effectiveness of RAI, external beam radiation, or other systemic therapies).

Furthermore, additional studies will be required to determine if there is any significant incremental benefit of adding these specific prognostic variables to the 2009 Initial Risk Stratification system.

Since the response to therapy dynamic (ongoing) risk stratification systems were primarily optimized and validated on DTC patients that had total thyroidectomy and RAI remnant ablation or adjuvant therapy, additional studies are needed to refine the definitions of excellent, biochemical incomplete, structural incomplete, and indeterminate responses in patients treated with total thyroidectomy without RAI ablation and in patients treated with less than total thyroidectomy (602). Furthermore, additional studies are needed to define what types of cross-sectional/functional imaging are required to rule out structural disease in order to classify a patient as having a biochemical incomplete response to therapy (based on initial risk, serum Tg, signs/symptoms, or other imaging results).

Another area of significant research interest centers on identifying specific clinical situations in which the molecular findings provide clinically meaningful information that goes beyond what is predicted by standard clinico-pathological staging. Molecular findings that provide these types of prognostic information or guide optimal initial/ongoing treatment decisions have the potential to significantly alter clinical management.

[D5] Improving our understanding of the risks and benefits of DTC treatments and optimal implementation/utilization

In achieving a better understanding of the risks and benefits of DTC treatments (such as extent of primary surgery or secondary surgery for recurrence, RAI ablation/treatment, and thyroid hormone suppressive therapy), more prospective long-term outcome research is needed, and in particular, RCTs. In the case of relatively uncommon adverse effects of treatments, prospective surveillance research is also needed. Vaisman and colleagues (633) are conducting a prospective, nonrandomized study to determine which patients who have low-to intermediate-risk DTC should receive RAI remnant ablation based on the postoperative stimulated serum Tg level. Longer-term follow-up and randomized multicenter studies are needed to determine how this commonly used biomarker can be applied to decision-making in a majority of patients with low- to intermediate-risk DTC. Prospective collection of data on quality of life and related outcomes (when relevant) is also needed in DTC trials. Using an evidence-based approach to knowledge synthesis (systematic reviews, meta-analyses, and clinical practice guidelines) of data can also enable evidence-informed clinical practice. Barriers to dissemination and implementation of clinical practice guideline recommendations need to be overcome. Evidence-based guidelines may need to be formally adapted to various practice settings to enable their implementation. Important gaps in patient-directed knowledge translation have been recently identified by thyroid cancer patients and survivors (1060,1061). The development of plain language educational materials, including decision aids or other decision support tools would be helpful for use as adjuncts in physician counseling of patients about diagnostic and treatment options. Decision aids and other decision support have been associated with improvement in patients’ medical knowledge and reducing decisional conflict, in general oncology (1062,1063), with a recent RCT demonstrating benefits of a decision aid in thyroid cancer survivors considering RAI remnant ablation (1064).

[D6] Issues with measurement of Tg and anti-Tg antibodies

Current methodologies for both Tg and anti-Tg antibodies remain problematic in many ways that hopefully will be overcome in the future. For Tg assays, these include interference with Tg measurement by the presence of anti-Tg and heterophile antibodies and the use of a host of different methods with varying results in terms of sensitivity or detection limits. Assay calibration or standardization may be successful only if the same certified reference standard is employed, which currently is CRM-457. While more “ultrasensitive” Tg assays have been developed, we need to determine the true clinical significance or utility of measurable levels below 0.2 ng/mL (either on suppression or after TSH stimulation) as indicating evidence of residual disease or outcome. Recently developed mass spectrometry–based assays have offered some promise, but are yet to be validated (806). Competitive immunoassays have not provided an alternative in view of their unpredictability (771). In regard to anti-Tg antibodies, we need to better characterize the various epitopes of interfering antibodies to better understand their effect in different sera in order to interpret the associated spectrum of results obtained for measurable Tg. We will need to do better than approximating Tg levels by the ratio of ICMA Tg to Tg measured by competitive immunoassay (1065). Authoritative bodies such as the National Academy of Clinical Biochemistry should consider mandating specific methodology rather than recommending general guidance. It remains to be seen whether the future standard might entail adoption of a modified ICMA, radioimmunoassay, or mass spectrometry methodology.


With the advent of improved technology, increased utilization, and specialized operator experience, US imaging can identify small-volume metastatic cervical lymph nodes. From the surgical pathology literature analyzing specimens
from prophylactic lateral and central neck dissections with normal preoperative cervical lymph node sonography, up to 90% of patients with papillary cancers <1 cm have metastatic level VI lymph nodes and up to 40% have metastatic lateral neck lymph nodes (359,1066,1067). Yet, in the absence of these dissections, this is not the observed clinical loco-regional recurrence rate for these patients. It should not be surprising then, that during extended surveillance, US will be able to detect small-volume metastatic disease that may represent a stable reservoir of residual cancer. On the other hand, grossly involved metastatic lymph nodes were at one time minimally invaded by metastatic thyroid cancer. The challenge is to differentiate between low-volume metastatic disease that progresses with potential clinical consequences, and that which remains stable. To date, only one study has addressed this question, and no sonographic, pathologic, demographic, or molecular feature predicted outcome (849). In addition, it is unclear if growth itself is a harbinger of decreased survival. Therefore, observational studies using standardized US scanning protocols are first required to define the magnitude of this scenario and define predictors of disease progression. Subsequently, randomized controlled intervention trials could be designed to address change in outcome, such as development of additional loco-regional disease, appearance of distant metastases, or disease-specific survival.

**[D8] Novel therapies for systemic RAI-refractory disease**

Over the years, a wide variety of agents have been used in preclinical (valproic acid, trichostatin, depsipeptide, 5-azacytidine, arsenic tri-oxide) and clinical (retinooids, thiazolidinediones) models to “redifferentiate” thyroid cancer cells in order to restore RAI avidity. While these agents showed limited clinical effectiveness, the observation that oncogenic activation of the MAP kinase pathway was associated with down-regulation of the genes involved in iodine metabolism suggested an alternative, targeted approach to redifferentiation therapy. Recently, two proof-of-principle pilot clinical trials have confirmed that targeted blockade of the MAP kinase pathway can result in clinically relevant restoration of RAI avidity in a substantial percentage of RAI-refractory thyroid cancer patients (1052,1068). In the first trial, 1 month of the MEK inhibitor selumetinib increased RAI uptake in 12 of 20 iodine-refractory thyroid cancer patients, with structural tumor shrinkage seen in five of the eight patients that achieved lesional dosimetry high enough to warrant RAI therapy (1052). In the second trial, 1 month of the BRAF inhibitor dabrafenib restored RAI avidity on diagnostic WBS in 6 of 10 RAI-refractory patients, resulting in structural responses in two patients and a decrease in serum Tg in four patients (1068). While the primary focus of current redifferentiation trials has been in the setting of RAI-refractory distant metastases, future studies are needed to define the role of redifferentiation therapy in the high-risk adjuvant therapy setting and in the RAI-responsive metastatic disease setting in an effort to enhance the tumoricidal effect of RAI before the tumors become RAI refractory.

Efforts to develop additional and further improved systemic therapeutic approaches to RAI-refractory metastatic DTC presently involve a wide a variety of approaches. First, the question arises as to whether kinase inhibitors with already demonstrated clinical activity in DTC can be made more effective as single agents. For example, based upon data indicating strong correlation between achieved pazopanib levels and extent of clinical response in DTC, indicating absence of response to pazopanib therapy in the lowest quintile of pazopanib drug levels, one clinical trial is presently examining the feasibility and potential benefits of individualization of pazopanib drug levels with the goal of achieving target drug levels in the highest achievable fraction of patients, in hopes of improving the fraction of patients benefiting (NCT01552356). Another study is examining the differential impacts of continuous versus intermittent pazopanib dosing in thyroid cancer (NCT01813136).

Generally predicated upon identification of synergistic interactions in preclinical models, several studies are examining the question of whether therapy combining several agents may improve outcomes in thyroid cancer. Although the majority of combinatorial studies assess effects of multiple coadministered small molecule therapeutics, several are using kinase inhibitors in combination with RAI in efforts to enhance RAI avidity and clinical efficacy in the context of RAI-refractory disease (e.g., NCT00970359). Still other trials are focusing on novel nontraditional therapies, including engineered candidate virotherapeutics (e.g., NCT01229865).

Alternatively, another active area of investigation involves efforts to therapeutically target specific alterations (mutations, translocation) found in thyroid cancers to individualize therapy and thereby potentially improve outcomes in the process; one such recent example is the use of the BRAFV600E inhibitor vemurafenib in BRAFV600E PTC (NCT01286753), with many additional mutation-specific therapeutic trials under development or underway.

Finally, immunotherapy including checkpoint inhibitors (e.g., PD-1/PD-L1) has shown promise in other cancers (1069–1071) and is being investigated in advanced RAI-refractory thyroid cancer.

**[D9] Survivorship care**

The American Cancer Society estimates that there will be 62,450 new cases of thyroid cancer diagnosed in 2015, but only 1950 deaths from thyroid cancer (www.cancer.org/cancer/thyroidcancer/detailedguide/thyroid-cancer-key-statistics). According to SEER database statistics, there are more than 600,000 people living with thyroid cancer in the United States alone (seer.cancer.gov/statfacts/html/thyro.html). Despite these large numbers of patients living with thyroid cancer, there is only a modest amount of peer-reviewed literature studying thyroid cancer survivors.

The majority of the literature involving thyroid cancer survivors relates directly to the short- and long-term effects of thyroid cancer therapies: surgery, RAI, and lifetime thyroid hormone therapy. There is very little information regarding the impact of the diagnosis itself or the effect of living with persistent disease such as Tg-positive, scan-negative thyroid cancer. A thoughtful and comprehensive analysis of thyroid cancer survivors will likely require both qualitative studies with in-depth interviews of survivors that will represent as many demographics of thyroid cancer patients as possible. Additionally, or as a next step, the development and/or utilization of a validated survey type instrument needs to be developed. This instrument would be designed to assess the quality of life of thyroid cancer survivors in a more quantitative manner, allowing for rigorous statistical analyses, and will help identifying areas to target that may improve the lives of thyroid cancer survivors (1072).
Further research also needs to be performed in addressing patient and survivor care needs throughout the active treatment and survivorship trajectory, including issues such as identification and management of treatment-related side effects (1060,1061), psychosocial distress (1060,1061), persistent fatigue (1073,1074), financial impact (12), and cancer-related worry (1075).

ACKNOWLEDGMENTS

The task force wishes to thank Ms. Bobbi Smith, Executive Director, ATA, and Ms. Sharleene Cano, Assistant to the Taskforce, for their constant help and support, as well as Ms. Vicki Wright (Division of Endocrinology, University of Colorado School of Medicine) for her assistance in manuscript preparation. We thank Dr. Joshua Klopper (Division of Endocrinology, University of Colorado School of Medicine) on behalf of the CAC Quality Of Life Task Force for the contribution of the Survivorship Care section [D9]. We also thank Dr. Irwin Klein (North Shore University Hospital, Manhasset, New York) for his input on TSH targets for long-term thyroid hormone therapy (Table 15). We would like to thank the ATA members who responded to our survey for preparation in for this iteration of the guidelines as well as manuscript review prior to journal submission, and the leadership of ThyCa: Thyroid Cancer Survivors’ Association, Inc. and Thyroid Cancer Canada who provided written feedback on our survey. A.M.S. holds a Cancer Care Ontario Health Services Research Chair, which enabled protected time for research and contribution to these guidelines. These guidelines were funded by the ATA without support from any commercial sources. The patient organization, ThyCa: Thyroid Cancer Survivors’ Association, Inc., contributed an unrestricted educational grant toward the development of the thyroid nodules and differentiated thyroid cancer guidelines.

The following groups reviewed and endorsed the final document: American Association of Clinical Endocrinologists; American Association of Endocrine Surgeons; American Head and Neck Society; Asia Oceania Thyroid Association (AOTA); British Society of Nuclear Medicine; Canadian Association of Otolaryngology Head and Neck Surgery; The Endocrine Society; Endocrine Society of Australia; European Thyroid Association; International Association of Endocrine Surgeons; International Federation of Head and Neck Oncologic Societies; Japanese Thyroid Association; Latin American Thyroid Society; Society of Surgical Oncology; Ukrainian Association of Endocrine Surgeons.

DISCLAIMER

It is our goal in formulating these guidelines, and the ATA’s goal in providing support for the development of these guidelines, that they assist in the clinical care of patients and share what we believe is current, rational, and optimal medical practice. In some circumstances, it may be apparent that the level of care recommended may be best provided in limited centers with specific expertise. Finally, it is not the intent of these guidelines to replace individual decision-making, the wishes of the patient or family, or clinical judgment.

AUTHOR DISCLOSURE STATEMENT

These guidelines were funded by the ATA without support from any commercial sources. K.C.B., G.M.D., F.P., G.A.R., A.M.S., and K.S. have no significant financial or competing interests to disclose.

B.R.H. has received grant/research support from Genzyme and Novo Nordisk. E.K.A. has received research support from Asuragen, Veracyte, and Genzyme. She has been an advisor for Genzyme. E.K.A. has received grant support from Asuragen and has been a CME speaker for Genzyme. Y.N. has been a consultant for Quest Diagnostics. His institution, UPMC, has received one-time speaker honorarium from Genzyme. E.K.A. and R.M.T. received consulting payments through stock options from Genzyme. E.K.A. has received research support from Asuragen, Genzyme, Novo Nordisk, and Eli Lilly. E.K.A. has received speaker honoraria from Genzyme, Bayer, AstraZeneca, Eli Lilly, and Bayer. E.K.A. has received grant/research support from Asuragen and has been a CME speaker for Genzyme. Y.N. has been a consultant for Quest Diagnostics. His institution, UPMC, has received one-time speaker honorarium from Genzyme and is a member of the International Thyroid Oncology Group. J.A.S. has received one-time speaker honorarium from Eli Lilly and is a member/ATA representative on the Data Monitoring Committee for the Medullary Thyroid Cancer Registry called for by the FDA and funded by NovoNordisk, Astra Zeneca, GlaxoSmithKline, and Eli Lilly. D.L.S. has received grant/research support from AstraZeneca. M.S. has received one-time speaker honorarium from Genzyme and is a consultant on behalf of the CAC Quality Of Life Task Force for the contribution of the Survivorship Care section [D9]. We also thank Dr. Joshua Klopper (Division of Endocrinology, University of Colorado School of Medicine) on behalf of the CAC Quality Of Life Task Force for the contribution of the Survivorship Care section [D9]. We also thank Dr. Irwin Klein (North Shore University Hospital, Manhasset, New York) for his input on TSH targets for long-term thyroid hormone therapy (Table 15). We would like to thank the ATA members who responded to our survey for preparation in for this iteration of the guidelines as well as manuscript review prior to journal submission, and the leadership of ThyCa: Thyroid Cancer Survivors’ Association, Inc. and Thyroid Cancer Canada who provided written feedback on our survey. A.M.S. holds a Cancer Care Ontario Health Services Research Chair, which enabled protected time for research and contribution to these guidelines. These guidelines were funded by the ATA without support from any commercial sources. The patient organization, ThyCa: Thyroid Cancer Survivors’ Association, Inc., contributed an unrestricted educational grant toward the development of the thyroid nodules and differentiated thyroid cancer guidelines.

The following groups reviewed and endorsed the final document: American Association of Clinical Endocrinologists; American Association of Endocrine Surgeons; American Head and Neck Society; Asia Oceania Thyroid Association (AOTA); British Society of Nuclear Medicine; Canadian Association of Otolaryngology Head and Neck Surgery; The Endocrine Society; Endocrine Society of Australia; European Thyroid Association; International Association of Endocrine Surgeons; International Federation of Head and Neck Oncologic Societies; Japanese Thyroid Association; Latin American Thyroid Society; Society of Surgical Oncology; Ukrainian Association of Endocrine Surgeons.

REFERENCES


38. Mazeh H, Benavidez J, Poehls JL, Youngwirth L, Chen H, Sippel RS 2012 In patients with thyroid cancer of follicular cell origin, a family history of nonmedullary thyroid cancer in one first-degree relative is associated with more aggressive disease. Thyroid 22:3–8.


40. Richards ML 2010 Familial syndromes associated with thyroid cancer in the era of personalized medicine. Thyroid 20:707–713.


90. Tae HJ, Lim DJ, Baek KH, Park WC, Lee YS, Choi JE, Lee JM, Kang MI, Cha BY, Son HY, Lee KW, Kang SK 2007 Diagnostic value of ultrasonography to distinguish between benign and malignant lesions in the management of thyroid nodules. Thyroid 17:461–466.
111. Redman R, Zalaiznck H, Mazzaferrin EL, Massoll NA 2006 The impact of assessing specimen adequacy and number of needle passes for fine-needle aspiration biopsy of thyroid nodules. Thyroid 16:55–60.


125. Deleted.


127. Deleted.


144. Chow SM, Law SC, Chan JK, Au SK, Yau S, Lau WH 2003 Papillary microcarcinoma of the thyroid—prog-
98 HAUGEN ET AL.


216. Smith RB, Robinson RA, Hoffman HT, Graham MM 2008 Preoperative FDG-PET imaging to assess the ma-


298. Boi F, Baghino G, Atzeni F, Lai ML, Faà G, Mariotti S 2006 The diagnostic value for differentiated thyroid carcinoma metastases of thyroglobulin (Tg) measurement in washout fluid from fine-needle aspiration biopsy of neck lymph nodes is maintained in the presence of circulating anti-Tg antibodies. J Clin Endocrinol Metab 91:1364–1369.


315. Padovani RP, Kasamatsu TS, Nakabashi CC, Camacho CP, Andreoni DM, Malouf EZ, Marone MM, Maciel RM, Bissopa RP 2012 One month is sufficient for urinary iodine to return to its baseline value after the use of water-soluble iodinated contrast agents in post-thyroidectomy patients requiring radioactive iodine therapy. Thyroid 22:925–930.


317. McLeod DS, Cooper DS, Ladenson PW, Ain KB, Brierley JD, Fein HG, Haugen BR, Jonklaas J, Magner J, Ross DS, Skarulis MC, Steward DL, Maxon HR, Sherman SI, for the National Thyroid Cancer Treatment Cooperative Study Group 2014 Prognosis of differentiated thyroid cancer in relation to serum thyrotrpin and thyroglobulin antibody status at time of diagnosis. Thyroid 24:35–42.


328. Vaisman F, Shaha A, Fish S, Michael TR 2011 Initial therapy with either thyroid lobectomy or total thyroidectomy without radioactive iodine remnant ablation is associated with very low rates of structural disease recurrence in properly selected patients with differentiated thyroid cancer. Clin Endocrinol (Oxf) 75:112–119.


335. Randolph GW, Duh QY, 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.


338. Lebouleux S, Rubino C, BAUdIN E, Caillou B, Hartl DM, Bidart JM, TravaglI JP, SchulumBerger M 2005 Prognostic factors for persistent or recurrent disease of papillary thyroid carcinoma with neck lymph node metastases and/or tumor extension beyond the thyroid capsule at initial diagnosis. J Clin Endocrinol Metab 90:5723–5729.


in patients with differentiated thyroid carcinoma: a randomized comparison between 1.1 and 3.7 GBq activities. Nucl Med Commun 34:767–770.


481. Koo JS, Hong S, Park CS 2009 Diffuse sclerosing variant is a major subtype of papillary thyroid carcinoma in the young. Thyroid 19:1225–1231.


505. Hirokawa M, Maekawa M, Kuma S, Miyauchi A 2010 Cribriform-morular variant of papillary thyroid carcino-
ma—cytological and immunocytochemical findings of 18 cases. Diagn Cytopathol 38:890–896.


541. Castagna MG, Maino F, Cipri C, Belardini V, Theodoropoulou A, Cevenini G, Pacini F 2011 Delayed risk stratification, to include the response to initial treatment (surgery and radioiodine ablation), has better outcome predictivity in differentiated thyroid cancer patients. Eur J Endocrinol 165:441–446.


556. Deleted.


ATA THYROID NODULE/DTC GUIDELINES


585. Deleted.


588. Soyluk O, Boztepe H, Aral F, Alagol F, Ozbey NC 2011 Papillary thyroid carcinoma patients assessed to be at
low or intermedietary risk after primary treatment are at greater risk of long term recurrence if they are thyro-
globulin antibody positive or do not have distinctively low thyroglobulin at initial assessment. Thyroid 21:1301–
1308.

vecchia G, Cabria M, Orlandi F 2010 Low thyroglobulin concentrations after thyroidectomy increase the prog-
nostic value of undetectable thyroglobulin levels on levo-thyroxine suppressive treatment in low-risk differ-

Limited value of repeat recombinant human thyrotropin (rhTSH)-stimulated thyroglobulin testing in differentiated
thyroid carcinoma patients with previous negative rhTSH-
stimulated thyroglobulin and undetectable basal serum

591. Kloos RT, Mazzaferr E 2005 A single recombinant
human thyrotropin-stimulated serum thyroglobulin mea-
surement predicts differentiated thyroid carcinoma me-
tastases three to five years later. J Clin Endocrinol Metab
90:5047–5057.

592. Kloos RT 2010 Thyroid cancer recurrence in patients
clinically free of disease with undetectable or very low
serum thyroglobulin values. J Clin Endocrinol Metab
95:5241–5248.

593. Han JM, Kim WB, Yim JH, Kim WB, Kim TJ, Ryu JS,
Gong G, Sung TY, Yoon JH, Hong SJ, Kim EY, Shong
YK 2012 Long-term clinical outcome of differentiated
thyroid cancer patients with undetectable stimulated
thyroglobulin level one year after initial treatment.
Thyroid 22:784–790.

594. Rosario PW, Furtado MS, Mineiro Filho AF, Lacerda
RX, Calsolari MR 2012 Value of repeat stimulated thy-
roglobulin testing in patients with differentiated thyroid
carcinoma considered to be free of disease in the first
year after ablation. Thyroid 22:482–486.

595. Grassard M, Borget I, Eder-Sanson A, Giraudet AL,
Mundler O, Tourbeau M, Bonichon F, Borson-Chazot F,
Leenhart L, Schwartz C, Dejax C, Brenot-Rossi I,
Toubert ME, Torlontano M, Benhamou E, Schlumberger
M 2011 Long-term follow-up of patients with papillary
and follicular thyroid cancer: a prospective study on 715

596. Pelletti H, Valimaki MJ, Loyttyinen E, Schalin-Iantt C
2010 Post-ablative serum thyroglobulin is an indepen-
dent predictor of recurrence in low-risk differentiated
thyroid carcinoma: a 16-year follow-up study. Eur J

597. Klubo-Giewezdinska I, Burman KD, Van Nostrand D,
Wartofsky L 2011 Does an undetectable rhTSH-stimulated
Tg level 12 months after initial treatment of thyroid
cancer indicate remission? Clin Endocrinol (Oxf) 74:111–
117.

598. Crocetti U, Durante C, Attard M, Maniglia A, Tumino S,
Bruno R, Bonifito N, Dicembrino F, Varraso A, Meringo
D, Filetti S, Trischitta V, Torlontano M 2008 Predictive
value of recombinant human TSH stimulation and neck ultrasonography in differentiated thyroid cancer
patients. Thyroid 18:1049–1053.

599. Torlontano M, Attard M, Crocetti U, Tumino S, Bruno
R, Costante G, D’Azzo G, Meringolo D, Ferretti E,
Sacco R, Arturi F, Filetti S 2004 Follow-up of low risk
patients with papillary thyroid cancer: role of neck ul-
trasonography in detecting lymph node metastases. J
Clin Endocrinol Metab 89:3402–3407.

600. Verburg FA, Stokkel MP, Duren C, Verkooijen RB,
Mader U, van Isselt JW, Marlowe RJ, Smit JW, Reiners
C, Luster M 2010 No survival difference after successful
(131)I ablation between patients with initially low-risk
and high-risk differentiated thyroid cancer. Eur J Nucl

601. Giovanella L, Maffioli M, Ceriani L, De PD, Spriano G
2009 Unstimulated high sensitive thyroglobulin mea-
surement predicts outcome of differentiated thyroid

602. Monesso DP, Tuttle RM 2014 Update on differentiated
thyroid cancer staging. Endocrinol Metab Clin North Am
43:401–421.

603. Lemb J, Hufner M, Meller B, Homayounfar K, Sahlmann
C, Meller J 2013 How reliable is secondary risk stratifi-
cation with stimulated thyroglobulin in patients with
differentiated thyroid carcinoma? Results from a retro-

604. Nascimento C, Borget I, Al Ghuzlan A, Deandreis D,
Chami L, Travagli JP, Hartl D, Lumbroso J, Chougnet C,
Lacroix L, Baudin E, Schlumberger M, Leboulleux S
2011 Persistent disease and recurrence in differentiated
thyroid cancer patients with undetectable postoperative
stimulated thyroglobulin level. Endocr Relat Cancer
18:R29–R40.

605. Webb RC, Howard RS, Stojadinovic A, Gaitonde DY,
Wallace MK, Ahmed J, Burch HB 2012 The utility of
serum thyroglobulin measurement at the time of
remnant ablation for predicting disease-free status in
patients with differentiated thyroid cancer: a meta-
analysis involving 3947 patients. J Clin Endocrinol Metab
97:2754–2763.

606. Chindris AM, Diehl NN, Crook JE, Fatourechi V,
Smallridge RC 2012 Undetectable sensitive serum thyroglobulin (0.1 ng/mL) in 163 patients with follicular cell-derived
tyroid cancer: results of rhTSH stimulation and neck ul-
trasonography and long-term biochemical and clinical

607. Vaisman F, Tal H, Grewal R, Tuttle RM 2011 In
differentiated thyroid cancer, an incomplete structural
response to therapy is associated with significantly worse
clinical outcomes than only an incomplete thyroglobulin

608. Verburg FA, Luster M, Cupini C, Chiovato L, Duntas,
Elisei R, Feldt-Rasmussen U, Rimmele H, Seregni E,
Smit JW, Theimer C, Giovannella L 2013 Implications of
thyroglobulin antibody positivity in patients with dif-
ferentiated thyroid cancer: a clinical position statement.
Thyroid 23:1211–1225.

609. Spencer CA, Takeuchi M, Kazarosyan M, Wang CC,
Gutler RB, Singer PA, Fatemi S, LoPresti JS, Nicoloff JT
1998 Serum thyroglobulin autoantibodies: prevalence, in-
fluence on serum thyroglobulin measurement, and prog-
nostic significance in patients with differentiated thyroid

610. Chung JK, Park YJ, Kim TY, So Y, Kim SK, Park DJ,
Lee DS, Lee MC, Cho BY 2002 Clinical significance of
elevated level of serum antithyroglobulin antibody in
patients with differentiated thyroid cancer after thyroid
ablation. Clin Endocrinol (Oxf) 57:215–221.
613. Adil A, Jafri RA, Waqar A, Abbasi SA, Matiul H, As- 
616. Thomas D, Liakos V, Vassiliou E, Hatzimarkou F, 
617. Castagna MG, Tala Jury HP, Cipri C, Belardini V, Fioravanti C, Pasqui L, Sestini F, Theodoropoulou A, 
618. Baudin E, Do Cao C, Cailleux AF, Lebouleux S, Tra- 
624. Padovani RP, Robenshtok E, Brokhin M, Tuttle RM 
625. Hsieh CJ, Wang PW 2014 Sequential changes of serum antithyroglobulin antibody levels are a good predictor of 
629. Rondeau G, Fish S, Hann LE, Fagin JA, Tuttle RM 2011 Ultrasonographically detected small thyroid bed nodules 
634. Valadao MM, Rosario PW, Borges MA, Costa GB, Re- 


661. Chen MK, Yasrebi M, Samii J, Staib LH, Doddamane I, Cheng DW 2012 The utility of I-123 pretherapy scan in


669. Verburg FA, Verkooijen RB, Stokkel MP, van Isselt JW 2009 The success of 131I ablation in thyroid cancer patients is significantly reduced after a diagnostic activity of 40 MBq 131I. Nuklearmedizin 48:138–142.


671. Jonklaas J, Sarlis NJ, Litofsky D, Ain KB, Bigos ST, Brierley JD, Cooper DS, Haugen BR, Ladenson PW, Magner J, Robbins J, Ross DS, Skarulis M, Maxon HR, Sherman SI 2006 Outcomes of patients with differentiated thyroid carcinoma following initial therapy. Thyroid 16:1229–1242.

672. Jonklaas J, Cooper DS, Ain KB, Bigos T, Brierley JD, Haugen BR, Ladenson PW, Magner J, Robbins J, Ross DS, Skarulis MC, Steward DL, Maxon HR, Sherman SI 2010 Radioiodine therapy in patients with stage I differentiated thyroid cancer. Thyroid 20:1423–1424.


697. Robbins RJ, Driedger A, Magnier J 2006 Recombinant human thyrotropin-assisted radioiodine therapy for patients with metastatic thyroid cancer who could not tolerate endogenous thyrotropin or be withdrawn from thyroxine. Thyroid 16:1121–1130.


707. Haugen ET AL.


712. Tala H, Robbins F, Fagin JA, Larson SM, Tuttle RM 2011 Five-year survival is similar in thyroid cancer patients with distant metastases prepared for radioactive iodine therapy with either thyroid hormone withdrawal or recombinant human TSH. J Clin Endocrinol Metab 96:2105–2111.


726. Sabra M, Grewal R, Ghossein RM, Tuttle RM 2014 Higher administered activities of radioactive iodine are associated with less structural persistent response in older, but not younger, papillary thyroid cancer patients with lateral neck lymph node metastases. Thyroid 24:1088–1095.


735. Souza Rosario PW, Barroso AL, Rezende LL, Padrao EL, Fagundes TA, Penna GC, Purisch S 2004 Post I-131


771. Spencer CA 2011 Clinical review: Clinical utility of thyroglobulin antibody (TgAb) measurements for patients with differentiated thyroid cancers (DTC). J Clin Endocrinol Metab 96:3615–3627.


786. Deleted.

may be avoided in thyroid cancer patients who have undetectable stimulated serum Tg levels after initial treatment. J Clin Endocrinol Metab 87:1499–1501.


820. Van Nostrand D, Moreau S, Bandaru VV, Atkins F, Chennupati S, Mete M, Burman K, Wartofsky L 2010 (124I) positron emission tomography versus (131I) planar imaging in the identification of residual thyroid tissue and/or metastasis in patients who have well-differentiated thyroid cancer. Thyroid 20:879–883.


877. Deleted.


923. Kulkarni K, Van Nostrand D, Atkins F, Aiken M, Burnquist L, Wartofsky L 2006 The relative frequency in which empiric dosages of radiiodine would potentially overtreat or undertreat patients who have metastatic well-differentiated thyroid cancer. Thyroid 16:1019–1023.


975. Chao M 2010 Management of differentiated thyroid cancer with rising thyroglobulin and negative diagnostic


1005. Wichers M, Benz E, Palmedo H, Biersack HJ, Grunwald, Mazzaferri EL 2002 Gonadal damage from 131I therapy.


1065. Spencer C, Petrovic I, Fatemi S 2011 Current thyroglobulin autoantibody (TgAb) assays often fail to detect interfering TgAb that can result in the reporting of falsely low/undetectable serum Tg IMA values for patients with differentiated thyroid cancer. J Clin Endocrinol Metab 96:1283–1291.


1078. Rosario PW, Mineiro Filho AF, Prates BS, Silva LC, Calsolari MR 2012 Postoperative stimulated thyroglobulin of less than 1 ng/mL as a criterion to spare low-risk patients with papillary thyroid cancer from radioactive iodine ablation. Thyroid 22:1140–1143.

Address correspondence to:

Bryan R. Haugen, MD
University of Colorado School of Medicine
Aurora, CO

E-mail: bryan.haugen@ucdenver.edu
Thyroid Nodules: Our Current Clinical Dilemma
What the radiologist needs to know, a review of current ATA criteria and TI-RADS

CASES

DR. GARY BRAHM
FRIDAY, OCTOBER 27, 2017
Case 1:

- Would you biopsy this nodule?
Case 2:

Would you biopsy this nodule?
Case 3:

- Would you biopsy this nodule?