Commentary

Thyroid Nodules 2006: Managing what has been known for over 50 years

Jeffrey R. Garber

Division of Endocrinology, Harvard Vanguard Medical Associates, Massachusetts, USA

The 50th anniversary of the publication of Mortensen, Woolner and Bennet’s paper entitled Gross and microscopic findings in clinically normal thyroid glands has just passed. Their paper reported that at autopsy at the Mayo Clinic approximately half of those not known to have thyroid disease pre-mortem had thyroid nodules (Table 1). Although we need to bear in mind that those who died of thyroid disease were excluded from the study, the vast majority of thyroid nodularity that was discovered at autopsy appeared to be clinically unimportant.

What have we learned since then? During the last five to ten years the answer to this question has been as important as, if not more important than, it has ever been since Mortensen’s landmark description. This is not simply due to the substantial increase in thyroid nodularity following exposure to head and neck radiation. There are additional reasons for this. In the USA, for example, physicians receive more of their training in ambulatory settings than they did several decades ago. Training programs, at least at some academic centers such as our own, presently devote more time to instructing trainees in the physical examination of the thyroid than my training program in the 1970s did. Imaging has also contributed. Carotid artery ultrasound, cervical Magnetic Resonance Imaging (MRI), and chest and neck Computed Tomography (CT), which are ordered for a myriad of reasons unrelated to the evaluation of the thyroid, regularly disclose “incidental” thyroid nodules.

However, the increasing application of thyroid ultrasound may be the major reason we are discovering more thyroid nodules. This is because additional nodules are regularly found when thyroid ultrasound is employed for the purpose of evaluating or aspirating a previously discovered nodule. Moreover, its use is likely to increase as a result of the American Association of Clinical Endocrinologists (AACE)/Associazione Medici Endocrinologi (AME) and the American Thyroid Association (ATA) newly published guidelines. ATA’s guideline states: “Thyroid sonography should be performed in all patients with one or more suspected thyroid nodules”. AACE/AME’s guideline also recommends that “All patients with palpable thyroid nodules or multinodular goiter should undergo ultrasound”, citing “to look for coincidental nodules” as a reason to do so. Thus, in a real sense sensitive ultrasound has enabled us and is being promoted, both intentionally and unintentionally, to “incidentally” detect thyroid nodularity as efficiently as Mortensen and his co-pathologists could by inspecting and sectioning thyroid glands at autopsy. As a result, we are confronting a previously silent epidemic of subclinical nodu-
lar thyroid disease that we have known about for a long time but were hoping, with good reason, to continue largely to ignore.

I will review some of what we have learned or re-learned since then. This overview will touch upon older diagnostic techniques and their refinements as well as some methods that may help us deal with our present dilemma, which can be framed as follows: How do we rationally deal with a growing epidemic of thyroid nodules while promoting ultrasound for diagnostic precision which, at the same time, increases the magnitude of the problem?

**PALPATION VS. SUBCLINICAL DISEASE**

Since 1955, additional autopsy, ultrasound, and surgical series have confirmed what Mortensen taught us, namely that subclinical disease is present about ten times as often as we suspect it clinically. Uncovering more subclinical nodular disease not only increases the number of nodules that we have to evaluate but, as studies have shown, causes the number of detected cancers to rise more or less proportionally. If we extrapolate from the Framingham Study data on the prevalence of thyroid nodules in adults, specifically 1.5% in males and 6.4% in females, we could potentially be confronted with a 15% prevalence of thyroid nodules in males and 64% in females. The latter, arguably, could become the largest epidemic known to man. However, comparing a pre-mortem prevalence of thyroid cancer of 0.25% to a prevalence in one autopsy series of 3.6% leads to an estimate that only about 1 in 15 cases of thyroid cancers are diagnosed pre-mortem. This is consistent with the following observation. Despite a substantial rise in the incidence of thyroid cancer in the USA, the overall death rate from thyroid cancer has declined, with the possible exception of older white males. This accords with what endocrinologists have long appreciated, namely that most thyroid cancers, even when not diagnosed, do not have an impact on survival, nor even probably on morbidity. Therefore, which population or subsets of populations might benefit from the detection of an enormous number of previously undetected thyroid nodules and cases of thyroid cancer? Posed in another way, what cost in human and economic terms would result from evaluating and treating many patients who would have been better off not worrying about thyroid cancer, undergoing surgery, or becoming dependent on thyroid hormone, calcium or pharmacologic doses of vitamin D?

**ALGORITHM**

Until relatively recently the limitations of physical examination of the thyroid gland could be viewed as a mixed blessing. The bad news was that we missed many thyroid nodules. The good news is that the vast majority of nodules that we missed were destined to remain asymptomatic and harmless.

Our current approach or “algorithm” (Figure 1) still starts with a history and physical examination. Patients should be questioned about exposure to ionizing radiation, family history of either medullary or papillary thyroid cancer, or familial syndromes such as Gardner’s syndrome, Carney Complex, and Cowden’s syndrome. Symptoms suggestive of rapid growth or obstruction should also be elicited. This is generally followed by a serum TSH determination. If the serum TSH value is low, isotopic imaging should follow, preferably with I (Io-
algorithm for palpable nodules groups “malignant” with “suspicious” while their incidentaloma algorithm omits the designation “suspicious” entirely, using “follicular neoplasia” for “indeterminate”. The ATA uses the term “indeterminate” to cover “suspicious, follicular lesion”, or “follicular neoplasm” as well as “suspicious” for “papillary carcinoma or Hurthle cell neoplasm”. Yet their algorithm separates “indeterminate” into “suspect for neoplasia” and “suspect for cancer”. I favor the Mayo Clinic’s approach which categorizes those aspirates with sufficient material to make a diagnosis as benign, malignant or suspicious. Within the suspicious or indeterminate group are those consistent with follicular neoplasm, Hurthle cell lesion, or suspicious for papillary thyroid carcinoma (PTC). The former two designations carry a risk of malignancy of approximately 15%, while the latter according to the Mayo and others carries a risk in excess of 50% of being malignant.

Most would proceed with a radioiodine scan to establish if a lesion consistent with a follicular neoplasm is autonomous, which essentially though not entirely rules out the possibility of malignancy. Although Hurthle lesions may be autonomous, 24 removal without scanning is generally recommended. This has become standard practice, even though there is no compelling basis for removing an autonomous Hurthle cell lesion. In the era when scans were done before FNA, a higher percentage of autonomous or “hot” nodules were never aspirated. No doubt some of these were undiagnosed, innocuous Hurthle cell adenomas.

In the case of the lesion that is suspicious for papillary thyroid carcinoma, virtually no data exist to support the role of isotopic imaging. Because, according to many authors, the odds are greater than a 50% chance that the lesion is malignant, removal is recommended.

We generally recommend an ipsilateral thyroidectomy in those cases presenting lesions consistent with follicular neoplasms that are not autonomous or Hurthle cell neoplasms. This is because the respective odds of malignancy are only about 15% each, making unilateral surgery, which does not cause profound hypothyroidism and avoids the risk

**Figure 1.** Algorithm of thyroid nodule management.
of permanent hypoparathyroidism and bilateral laryngeal nerve injury resulting from a bilateral procedure, sufficient in 85% of cases. On the other hand, in cases of lesions suspicious for papillary thyroid carcinoma, bilateral surgery is the preferred strategy for most. This is because the odds are greater than 50% that it will be necessary once a final diagnosis is made. In any event, in my opinion it is critical for physicians to discuss the pros and cons of unilateral vs. bilateral thyroidectomy with patients. These discussions should be based on whether or not there are contralateral nodules, the risk they carry for malignancy, whether or not they are autonomous—which is more common in areas of marginal iodine sufficiency and increases the risk for developing thyrotoxicosis over time—, and patient preference.

**SPECIFIC ISSUES**

**Calcitonin**

Some recommend calcitonin measurements as part of the initial evaluation of thyroid nodules.\(^{25}\) This approach has not been embraced in the USA for several reasons. It is costly\(^ {20}\) and yields false positives due to C cell hyperplasia, which has only been shown to be pre-malignant in hereditary forms of medullary thyroid carcinoma. Some authors point out that there is little evidence that incidentally discovered small carcinomas become clinically significant,\(^ {26}\) though others entertain a different view.\(^ {27}\) While stimulation with pentagastrin, to help determine whether elevated calcitonin levels are malignant in origin, is not presently available in the USA, when clinically warranted, basal calcitonin levels or calcium stimulated values should be determined.

**Hurthle cells present, but not clearly a Hurthle Cell neoplasm:**

Hurthle cells are ubiquitous. They are found in a spectrum of thyroid disorders, including the very common conditions of Hashimoto’s thyroiditis and multinodular goiter. How then does one distinguish a Hurthle cell lesion from a non-Hurthle cell lesion when a subpopulation of Hurthle cells is seen on an FNA specimen? One of the benefits of ultrasound guided FNA is accurate localization of the tip of the aspirating needle. This enables the operator to be much more certain that a nodule rather than tissue surrounding a nodule is being sampled. Under these circumstances a mixed population of cells consistent with Hashimoto’s thyroiditis in combination with Hurthle cells would be most consistent with Hashimoto’s thyroiditis. However, aspirates are often pooled before being processed and read. Moreover, a Hurthle cell neoplasm may co-exist with Hashimoto’s thyroiditis, particularly in instances when a discrete nodule is present. We sometimes address this by repeating the aspirate under ultrasound guidance, avoiding the use of suction or being careful to discontinue it before the needle is fully withdrawn from the nodule. This minimizes the chance for “contamination” with cellular material from surrounding tissue. Passes are not pooled and are read separately. If all aspirates reveal mixed populations of cells without a predominant population of Hurthle cells, the odds are low that one is dealing with a Hurthle cell neoplasm. However, this does not rule out the possibility that a Hurthle cell lesion is present in a gland with Hashimoto’s thyroiditis or vice versa. In cases that are uncertain or clinical and ultrasound assessment suggests that a neoplasm is present, surgery is recommended.

**The benign aspirate: Suppression Therapy**

**Basis and origins**

Employing thyroid hormone to suppress TSH levels is supported by a number of experimental and clinical observations. Over 50 years ago Maloof showed that hypophysectomy prevented radiation induced thyroid tumors in rats.\(^ {28}\) For 30 years more TSH binding has been found on the surface of benign rather than malignant thyroid lesions.\(^ {29}\) More recently, around the time the utility of suppression therapy began to be called into question, Fogelfeld reported that irradiated patients treated with suppression therapy after subtotal thyroidectomy developed fewer benign (not malignant) thyroid nodules than those who were not treated with thyroid hormone.\(^ {30}\)

**Does it work?**

Some studies conclude that there is a benefit from thyroid hormone suppression therapy for thyroid nodules. Most do not. A recent metanalysis of se-
ries of patients with solid nodules with benign FNAs followed on relatively aggressive suppression for at least 6 months also concludes that there is no clear-cut benefit. However, the authors indicate that there is a “trend” towards benefit, thus implying that it might be evident if the combined series included more subjects. Even those who favor consideration of suppression therapy concede that at most 30% respond—where response includes stabilization or regression—and that suppression therapy carries cardiac and bone risk, among others, resulting from thyroid hormone excess. The risk is not trivial since thyroid hormone excess is a frequent consequence of thyroxine therapy, even if TSH suppression is not a goal of therapy. The Colorado Thyroid Disease Prevalence Study, for example, determined that more that one-fifth of those taking thyroid medication, regardless of indication, were hyperthyroid or subclinically hyperthyroid.

Who are the best candidates for suppression therapy?

Thyroxine therapy still may play a role in young, healthy patients, with relatively small thyroid nodules or those with nodular goiters without evidence of autonomy, particularly in regions with marginal iodine sufficiency.

Why employ suppression therapy?

Alexander’s ultrasound study demonstrated that most (89%) benign nodules grow by at least 15% over a period of 5 years. This degree of clinically inapparent growth approximates the upper limit of ultrasound exam to check variation reported by some, but falls far below the 50% change in nodule volume that Brauer and colleagues caution is required by ultrasound measurements before concluding that a nodule has truly grown. In any event, 22% of Alexander’s cohort who were followed up were re-biopsied. Those who were re-biopsied had a mean increase in nodule volume of 69%. Only one of the 74 patients who were re-biopsied was found to have a malignancy.

As more and more thyroid nodules, particularly ones that are not readily palpable, are monitored with ultrasound examinations, we will more frequently need to face the question of what to do with a patient with a previously benign aspirate whose nodule has grown. Most will be re-biopsied. While physicians tend to minimize the consequence and cost of repeating FNAs, many patients do not. After all, in the face of growth the reason for doing so is to address the anxiety provoking possibility that a malignancy has been missed. More importantly, many, despite benign FNAs, will undergo surgery in keeping with guideline recommendations.

Based on the above, properly selected patients could benefit from suppression therapy by having fewer FNAs and less surgery for benign disease. Although optimal TSH targets for suppression therapy are not well established, it “should never be fully suppressive” and has to be taken on a long-term basis.

The inadequate aspirate:

When strict criteria are employed for determining specimen adequacy or sufficiency, a substantial number of FNAs will be “inadequate” or yield insufficient material to make a diagnosis. The number of cells or clusters of cells deemed adequate for making a diagnosis varies, ranging from 5 to 6 groups of 10 or more cells to 10 groups of 20 cells, with one author recommending that there be at least 6 groups on two separate aspirates. Whether an adequate number of cells is necessary in some circumstances, such as the specimen rich in colloid, is also the subject of some debate. When an FNA is deemed inadequate, there is general agreement that it should be repeated. Even those who do not endorse using ultrasound for all initial aspirations generally favor doing so after an inadequate specimen is obtained. In one series, the success rate of employing ultrasound for a second aspirate ranged from 76% to 85% for solid nodules and those less than 25% cystic to 48% for nodules that were more than 75% cystic. What should be done after a second aspirate proves to be inadequate is less straightforward. Some will do a third because it is successful up to 30% of the time. We will do an US guided FNA a third time on portions of the nodule shown by Doppler to have blood flow. This may be particularly useful when dealing with a semi-cystic nodule. We believe, but have not proven, that solid peripheral regions of preserved flow are more likely
to yield adequate specimens than areas with decreased flow. A similar recommendation for the use of Doppler has recently been advocated by others.\(^{14}\) This is because deeper, solid appearing areas may be quite fibrotic and contain organized thrombus—with relatively few intact follicular cells amenable to FNA. After a third failure\(^{42}\) most will send the patient to surgery. We individualize this recommendation and may, after discussing this with the patient, recommend observation if the nodule is relatively small (less than 1.5 cm) and principally cystic. This is supported by an unpublished observation on a cohort of 778 patients at the Brigham and Women’s Hospital. None of 7 purely cystic lesions proved to be malignant and only 2 of 87 lesions that were 75% or more cystic were malignant (Frates MC, personal communication). On occasion, thyroxine suppression is offered to patients with principally semi-cystic lesions and inadequate aspirations, but virtually no data exist to support using it in this setting.

THE INDETERMINATE ASPIRATE: WHAT IS NEW AND WHAT IS ON THE HORIZON?

Current approaches to suspicious or indeterminate aspirates have been reviewed above. I will touch on some potentially useful new tools.

An interesting, though presently economically prohibitive, approach is the predictive value of PET scan negativity in this group. Until recently, most PET scan reports dealing with nodular thyroid disease focused on the high incidence of malignancy, up to 50% in some series,\(^ {43}\) in patients with PET positive thyroid incidentalomas. In a prospective series done at the Beth Israel Deaconess Medical Center, 31 patients with 48 lesions going to surgery regardless of their FNA findings had pre-operative PET scans. Twenty-one of 22 patients with indeterminate FNAs that were PET negative (SUV <5) had benign pathology, yielding a 95% negative predictive value.\(^ {44}\) A more recent study by de Geus-Oei yielded a similar result.\(^ {45}\)

If FDG thyroid imaging could be done economically, perhaps by limiting images to the thyroid region, and these results were confirmed in other studies, a negative PET study on an indeterminate aspirate might be a useful tool, particularly in a person who is at high risk for surgery.

What about the utility of markers? Some single markers that initially seemed quite promising have not panned out. These include galectin 3, PAX8-PPAR-gamma, and telomerase. Aspirates negative for human bone marrow endothelial cell (HBME-1) are highly unlikely to be papillary cancers, but cannot distinguish follicular carcinoma from follicular adenoma, since either can be positive or negative.\(^ {46}\) RET/PTC classification reports have yielded conflicting results.\(^ {47,48}\)

Detection of BRAF mutations presently appears promising with one series showing no false positive results for papillary thyroid cancer and the ability to diagnose papillary thyroid cancer in 5 of 15 indeterminate aspirates and one inadequate specimen.\(^ {49}\) Although this is promising, marker studies will ultimately have to be confirmed in several prospective series employing techniques suitable for pre-operative FNA specimens.

CONCLUSION:

As we pass the 50th anniversary of Mortensen’s 1955 report and are well into the “Ultrasound Era”, we are being confronted with the issue of how to best manage a longstanding but previously under-recognized “epidemic” of subclinical and clinical nodular thyroid disease. We must continue to strive to increase the net benefit to society that is gleaned from evaluating and treating thyroid nodules. Doing so will require endocrinologists, among others, to:

- Keep diagnostic and surveillance costs in check;
- Refine criteria for distinguishing nodules that require a definitive diagnosis from those that do not;
- Develop more accurate diagnostic techniques including applicable and reliable “molecular tools”\(^ {50}\);
- Use emerging guidelines to assist rather than restrict us in our approach to individual patients.

Hopefully, well before the 100th anniversary of Mortensen’s report, we will be able to claim more than a modicum of success in achieving these goals.
REFERENCES

3. Cooper DS, Doherty GM, Haugen BR, et al, Management Guidelines for Patients with Thyroid Nodules and Differentiated Thyroid Cancer. Thyroid 16: 1-33.
9. Blum M, Hussain MA, 2003 Evidence and thoughts about thyroid nodules that grow after they have been identified as benign by aspiration cytology. Thyroid 13: 637-641.
50. Davies TF, 2006 Is consensus a good thing in the management of thyroid nodules? Thyroid 16: 205.