Meltzer Charles (Orcid ID: 0000-0002-5216-9751)

AHNS Endocrine Section Clinical Consensus Statement:

North American Quality Statements and Evidence-Based Multidisciplinary Workflow Algorithms for the Evaluation and Management of Thyroid Nodules

Charles J. Meltzer, MD,[1] Jonathan Irish, MD, MSc,[2] Peter Angelos, MD, PhD[3], Naifa L. Busaidy, MD,[4] Louise Davies, MD, MS,[5] Sunshine Dwojak, MD, MPH,[6] Robert L. Ferris, MD, PhD,[7] Bryan R. Haugen, MD,[8] R. Mack Harrell, MD,[9] Megan R. Haymart, MD,[10] Bryan McIver, MD, PhD,[11] Jeffrey I. Mechanick, MD,[12] Eric Monteiro, MD, MSc,[13] John C. Morris, III, MD,[14] Luc G. T. Morris, MD, MSc,[15] Michael Odell, MD,[16] Joseph Scharpf, MD,[17] Ashok Shaha, MD,[18] Jennifer J. Shin, MD, SM,[19] David C. Shonka, Jr., MD,[20] Geoffrey B. Thompson, MD,[21] R. Michael Tuttle, MD,[22] Mark L. Urken, MD,[23] Sam M. Wiseman, MD,[24] Richard J. Wong, MD,[25] Gregory Randolph, MD[26]

[1] Kaiser Permanente Northern California, Santa Rosa, California

[2] Department of Otolaryngology, Head and Neck Surgery/Surgical Oncology, Princess Margaret Cancer Centre, University Health Network/University of Toronto, Toronto, Ontario, Canada

[3] Department of Surgery and MacLean Center for Clinical Ethics, The University of Chicago Medicine, Chicago, Illinois

[4] Department of Endocrine Neoplasia & Hormonal Disorders, The University of Texas MD Anderson Cancer Center, Houston, Texas

[5] The VA Outcomes Group, Department of Veterans Affairs Medical Center, White River Junction, Vermont; Section of Otolaryngology, Geisel School of Medicine at Dartmouth, Hanover, New Hampshire; The Dartmouth Institute for Health Policy and Clinical Practice, Hanover, New Hampshire

[6] Northwest Permanente Medicine, Kaiser Permanente Northwest, Portland, Oregon

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/hed.25526

[7] Department of Otolaryngology, Division of Head and Neck Surgery, University of Pittsburgh Cancer Institute, Pittsburgh, Pennsylvania

[8] Division of Endocrinology, University of Colorado School of Medicine, Aurora, Colorado

[9] Memorial Healthcare System—Departments of Integrative Endocrine Surgery and Pathology, Hollywood, Florida

[10] Division of Metabolism, Endocrinology, and Diabetes, Department of Internal Medicine, University of Michigan, Ann Arbor, Michigan

[11] Department of Biostatistics and Bioinformatics, H. Lee Moffitt Cancer Center and Research Institute, Tampa, Florida

[12] Divisions of Cardiology and Endocrinology, Diabetes and Bone Disease, Icahn School of Medicine at Mount Sinai, Mount Sinai School of Medicine, New York, New York

[13] Department of Otolaryngology—Head and Neck Surgery, Mount Sinai Hospital, University of Toronto, Toronto, Ontario, Canada

[14] Division of Endocrinology, Diabetes, Metabolism, and Nutrition, Mayo Clinic, Rochester, Minnesota

[15] Department of Surgery, Memorial Sloan Kettering Cancer Center, New York, New York

[16] Department of Otolaryngology—Head and Neck Surgery, The Ottawa Hospital, University of Ottawa, Ottawa, Ontario, Canada

[17] Department of Otolaryngology—Head and Neck Surgery, Head and Neck Institute, Cleveland Clinic, Cleveland, Ohio

[18] Department of Surgery, Head and Neck Service, Memorial Sloan Kettering Cancer Center, New York, New York

[19] Department of Otolaryngology, Harvard Medical School, Boston, Massachusetts

[20] Department of Otolaryngology—Head and Neck Surgery, University of Virginia Health System, Charlottesville, Virginia

[21] Departments of Surgery, Mayo Clinic, Rochester, Minnesota

[22] Department of Medicine, Endocrinology Service, Memorial Sloan Kettering Cancer Center, New York, New York

[23] Department of Otolaryngology Head and Neck Surgery, Mount Sinai Beth Israel, Icahn School of Medicine at Mount Sinai, New York, New York; THANC (Thyroid, Head and Neck Cancer) Foundation, New York, New York

[24] Department of Surgery, St. Paul's Hospital & University of British Columbia, Vancouver, British Columbia, Canada

[25] Department of Surgery, Memorial Sloan Kettering Cancer Center, New York, New York

[26] Department of Otolaryngology, Massachusetts General Hospital, Boston, Massachusetts

Correspondence to:

Charles J. Meltzer, MD

The Permanente Medical Group

401 Bicentennial Way

Santa Rosa, CA 95403

Phone: 707.393.3262

Email: charles.meltzer@kp.org

4338 words, 1 table, 3 figures, 69 references

Running head: AHNS Quality statements and multidisciplinary workflow algorithms

Funding source: None

ABSTRACT

Background: Care for patients with thyroid nodules is complex and multidisciplinary, and research demonstrates variation in care. The objective was to develop clinical guidelines and quality metrics to reduce unwarranted variation and improve quality.

Methods: Multidisciplinary expert consensus and modified Delphi approach. Source documents were workflow algorithms from Kaiser Permanente Northern California and Cancer Care of Ontario based on the 2015 American Thyroid Association (ATA) management guidelines for adult patients with thyroid nodules and differentiated thyroid cancer.

Results: A consensus-based, unified pre-, peri-, and postoperative workflow was developed for North American use. Twenty-one panelists achieved consensus on 16 statements about workflow-embedded process and outcomes metrics addressing safety, access, appropriateness, efficiency, effectiveness, and patient centeredness of care.

Conclusion: A panel of Canadian and U.S. experts achieved consensus on workflows and quality metric statements to help reduce unwarranted variation in care, improving overall quality of care for patients diagnosed with thyroid nodules.

INTRODUCTION

Palpable thyroid nodules occur in 5% of women and 1% of men in iodine-sufficient areas.[1,2] Their prevalence varies with detection mode and increases with age —2-6% by palpation versus 19-68% by ultrasound.[3,4] Among all thyroid nodules, 7-15% are malignant, and the proportion varies with age, gender, radiation exposure, family history, and other factors.[5-10] At least 94% of thyroid carcinomas are differentiated thyroid cancer (DTC), primarily papillary thyroid carcinoma (PTC) and follicular thyroid carcinoma.[11,12] The annual incidence of thyroid cancer per 100,000 person years has increased from 4.6 in 1974-1977 to 14.4 in 2010-2013.[13] In the United States, 64,300 cases were predicted to be diagnosed in 2016.[14] This figure is largely attributable to incidentally detected subclinical disease.[15-17] Hence, by 2019, PTC is estimated to become the third most common cancer among U.S. women and has an associated annual cost between $18 and $21 billion dollars.[12] Thyroid nodule evaluation and treatment varies substantially in the U.S., with thyroidectomy rates ranging from 29 to 139 per 100,000 Medicare beneficiaries. The patterns of heterogeneity in thyroid care appear unrelated to healthcare access, physician supply, or socioeconomic status.[18]

Reducing unwarranted variation in disease management is pivotal for optimizing the quality of care for individuals diagnosed with thyroid nodules. Clinical practice guidelines are one way to achieve this goal, providing guidance that helps health care professionals and patients make appropriate healthcare decisions together. Clinical guideline development is accomplished by synthesizing clinical evidence, but their creation is not always carried out with the end user in mind, which can make them difficult to use.[19,20] Clinical algorithms, which are tools created from guidelines, can enhance implementation by organizing complex narratives and decision trees into a format that allows easier use in everyday practice.[21] In 2015, the American Thyroid Association (ATA) released revised management guidelines for adult patients with thyroid nodules and DTC.[5] The ATA guidelines inspired independent simultaneous initiatives at Kaiser Permanente Northern California (KPNC) and Cancer Care of Ontario (CCO) to codify them into algorithms.

Here, we outline the work and results of our cross-national harmonization through a multidisciplinary consensus process with representatives from key professional medical organizations involved in thyroid care in the U.S. and Canada. The resulting unified workflow algorithms are designed to optimize implementation of the ATA guidelines and for standardization of care. We selected process and outcomes metrics through a modified Delphi process that correspond to the workflow, as well as other key aspects of thyroid perioperative care. These metrics can be used to support measurement of adherence. The objective is to reduce unwarranted variations in care for patients diagnosed with thyroid nodules.

MATERIALS AND METHODS

Harmonizing existing workflow algorithms

In the first phase, a subgroup of authors (C.M., J.I., and G.R.) compared the contents of the Thyroid Cancer Diagnosis Pathway Map designed by CCO to the KPNC workflows. Although the formatting was different, the content was very similar in both algorithms and both included a color guide to assign responsibility by specialty for each step along the care pathway. The more compact KPNC format was selected as the basis for the remaining consensus building process.

In the second harmonization phase, the workflow algorithms were shared with the entire author panel and reviewed over a series of meetings to gather input and create consensus regarding the content. In the third and final phase, each figure was reviewed by a subgroup of authors responsible for the associated content for the purpose of creating narrative summaries to support the quality metric statements. After additional editing, each of the three figures was brought back to the entire author panel for approval.

Development of quality statements - modified Delphi method

Selection of panel members

Panel members were selected to represent surgeons and endocrinologists from North America and major professional societies, including the American Head and Neck Society (AHNS), American Academy of Otolaryngology-Head and Neck Surgery, American Association of Endocrine Surgeons, American Thyroid Association, American Association of Clinical Endocrinologists (AACE), and American College of Surgeons.

Statement selection and related evidence

A subgroup of 10 co-authors (B.M., C.M., E.M., G.R., J.M., J.J.S., J.I., L.D., M.O., L.G.T.M.) identified topics for statements that would undergo consensus development, taking into consideration existing and planned quality metrics at KPNC and COO.[22,23] Twelve metrics common to the two systems were included, along with five others (antibiotic use, 30-day mortality rates, postoperative voice evaluation, preoperative calcium level, outpatient surgery) that were unique to KPNC. The subgroup had considered an abbreviated list of metrics but included all 17 to enhance tailoring to local care settings. After discussion among panel members, the quality metric for outpatient surgery was dropped because the chosen metric of length of stay incorporated same-day surgery. Based on the proposed metrics, candidate statements were proposed in alignment with the Institute of Medicine's principles of health care quality, with the goal of addressing safety, access, appropriateness, efficiency, effectiveness, and patient centeredness.

Candidate statements and a concise summary of associated evidence from the relevant guideline were distributed to panel members. The following evidence appraisal systems had been used during ATA guideline development (the source document for the evidence).[5] For therapeutic interventions, the task force had used the American College of Physicians Grading

System for the appraisal of evidence on therapeutic interventions.[24] For appraisal of evidence on diagnostic tests, prognostic tests and risk stratification systems, the task force had devised a system based on the STARD, GRADE, QUADAS-2, and other scales.[25-29] These systems were combined with expert opinion reached by consensus if there were insufficient data to guide recommendations.

Definitions and refinements for the proposed statements were drawn from guidelines produced by the AACE and American College of Endocrinology clinical review of postoperative hypoparathyroidism[30] and KPNC and CCO documents. The AACE guidelines used for our process had also been developed through established protocols and assessments of evidence to support guideline development.[31-34]

Consensus process

Consensus was assessed in two rounds. In the first round, the entire expert panel rated 16 statements. The first author sent by email to panel members the link to an online survey (www.surveymonkey.com, San Mateo, California) containing candidate statements and brief instructions. Members of the expert panel reviewed proposed statements and independently rated their level of agreement or disagreement with each, using a Likert scale ranging from 1 ("Strongly Disagree") to 9 ("Strongly Agree"), with 5 representing a "Neutral" response. Panelists could also provide free-text comments for further discussion, which were logged. Survey results were tracked, summarized, and interpreted at each stage.

During a subsequent teleconference, panel members discussed the results, focusing on statements that had engendered comments, a desire for more detail, or suggestions for refinement. Discussion revealed varying interpretations of and suggestions for rewording 10 statements. Consequently, a subgroup (C.M., L.D., J.S., J.I., G.R.) modified 10 statements for review in the second round (Table 1), clarifying areas prone to variable interpretation and ensuring statements were suitable for most practice settings. Six statements were neither revised nor reviewed again. The results were shared in a second telephone conference, and the final list of statements was sorted into preoperative, perioperative, and postoperative periods.

The Rand/UCLA Appropriateness method,[35,36] a modified Delphi method, [37,38] was used to quantitatively assess for panel consensus. A similar approach has been used in expert panels in other fields.[39,40] The consensus criteria required a median score $\geq$ 7.0, representing a response of “Agree” on the Likert scale. In addition, a disagreement index (DI) was calculated, defined as (interpercentile range/ interpercentile range adjusted for symmetry), with a value > 1.0 indicating disagreement.[35,40] Calculations and data analysis were performed by a Kaiser Permanente analytic team using Microsoft Excel 2016 (Redmond, Washington).

RESULTS

Workflow algorithm harmonization

The workflow algorithms were modified to include the American College of Radiology Thyroid Imaging Reporting and Data System (TIRADS), in addition to the ATA pattern recognition risk stratification system.[41] Minor revisions were made to the content within specific nodes of the figure by the leads responsible for each section (R.M.H., M.O., R.W.), along with visual changes to enhance the presentation. The resulting workflow algorithms address preoperative (Figure 1), perioperative (Figure 2), and postoperative (Figure 3) management of patients diagnosed with thyroid nodules.

Modified Delphi results

Twenty-one panelists completed both rounds of surveys. In the first round, consensus was achieved on 14 statements. At the completion of the second round, consensus was achieved on 16 statements; two statements that did not reach consensus in the first round did so in the second. Table 1 provides the results for both rounds of panel ratings.

Preoperative management (Figure 1)

The four essentials of preoperative thyroid nodule care are: 1) appropriate selection of patients for evaluation, 2) comprehensive anterior and lateral neck imaging with ultrasound, 3) careful clinical decision-making in accordance with existing guidelines (who to biopsy, how to interpret cytology /molecular testing and how to incorporate clinical history, imaging and laboratory studies into surgical planning) and (4) selection of a competent thyroid cancer surgeon.

Statement 1. Monitoring access to specialist consultation and elective surgery may help determine whether clinical care is readily available to patients. For patients diagnosed with euthyroid nodules, access to specialist consultation and elective surgery varies widely between the U.S. and Canada.[42] Patients waiting for thyroid surgery for lengthy periods of time may experience unnecessary stress and anxiety.[43] The panel did not make a recommendation about the upper limit of wait times for elective surgery. This statement supports setting-specific and locally relevant definitions of delayed access and improvement initiatives to reduce delays.

Statement 2. Risk stratification utilizing thyroid sonography with survey of the cervical lymph nodes should be performed in all patients with known or suspected thyroid nodules. The purpose of ultrasound-based risk stratification is to distinguish between thyroid nodules with the greatest and least potential for malignancy, which drives the need for either biopsy or observation. Appropriate ultrasound imaging is the foundation of risk stratification.[5] It is indicated in the presence of palpable thyroid nodules or the incidental discovery of thyroid pathology by unrelated imaging studies. Risk factors, such as a history of childhood or adolescent radiation exposure, a family history of thyroid cancer in multiple first-degree relatives, or certain clinical syndromes that have associated thyroid pathology, may drive decision making around fine needle aspiration (FNA) biopsy of a lesion that is smaller than the size described by the ATA guidelines.

When criteria for neck imaging are met, state-of-the-art thyroid nodule care requires ultrasound of the entire anterior and lateral neck, including: 1) central and lateral lymph nodes, 2) pyramidal lobes, 3) thyroglossal duct tract elements, and 4) relationships between thyroid nodules and surrounding strap muscles, vascular structures, trachea/esophagus, and posterior structures (bone, muscle, and recurrent laryngeal nerve [RLN]).[44]

The ATA thyroid nodule risk stratification[5] and radiologic TIRADS [41] systems incorporate clinical and imaging information to provide guidance for clinicians as to which thyroid nodules require observation, biopsy by FNA, or neither. Figure 1 combines the ATA 2015 and TIRADS 2017 classifications in the second and third row of cells: 1) ATA benign and TIRADS 1 and 2 nodules may be observed, 2) ATA very low to high suspicion and TIRADS 3-5 nodules should undergo FNA depending on nodule size, with biopsy cut points differing between the two systems, and 3) all ATA very low to high cancer suspicion and TIRADS 3-5 nodules below FNA size cut points may be observed with close follow up or biopsied, based on patient preference and risk factors.

Cross-sectional imaging using computed tomography (CT) with contrast or magnetic resonance imaging (MRI) should be considered when planning surgery if there are abnormal nodes at the limits of the sonogram or extensive nodal disease or the primary tumor is very large or invasive.

Statement 3. The institutional or system wide annual rate of thyroidectomies performed for DTC in a nodule with a maximum diameter less than 1 cm. should be monitored. The indiscriminate use of high-resolution ultrasonography has led to the diagnosis of many < 1 cm. non-palpable thyroid cancers, the vast majority of which would not otherwise require clinical attention.[45] The AHNS aims to minimize the human and financial costs of unnecessary evaluation and surgery by discouraging ultrasound imaging in patients without risk factors for thyroid cancer and normal physical examinations.

Statement 4. Utilization of Bethesda classification is important for consistent reporting of thyroid cytopathology. The Bethesda System for Reporting Thyroid Cytopathology[46] provides a diagnostic cytopathologic framework for interpreting and reporting thyroid cancer risk based on FNA of thyroid nodules.

For nodules qualifying for neither immediate surgery nor dismissal, ongoing follow-up with clinical exam or ultrasound may be performed at 6-month to 2- to 3-year intervals, depending on clinical context, and possible repeat FNA at 3 months after initial aspiration for Bethesda I or longer, depending on ultrasound findings. [47]

Statement 5. To avoid the potential complications associated with reoperations, it is suggested to obtain a parathyroid hormone or serum calcium level prior to thyroid surgery. As intraoperative parathyroid hormone (ioPTH) levels are being used to obtain a baseline at the time of a total thyroidectomy and/or at the conclusion of surgery to better direct the management of potential postoperative hypoparathyroidism,[48] an abnormally and unexpectedly high ioPTH level will raise the question of an incidental parathyroid adenoma. Incidental parathyroid adenomas that are diagnosed at the time of thyroid surgery are reported to occur in 0.2-4.5% of cases.[49] Obtaining a parathyroid hormone level or a serum calcium level should be considered prior to surgery as a useful screening lab to reduce the unexpected incidentaloma finding or need to re-operate on a patient who may have concomitant disease. Guidelines for appropriate management of incidentally identified parathyroid adenomas should be closely followed.

Statement 6. Surgical volume is a positive indicator of more efficient and effective care. Numerous studies document associations between higher surgeon volumes for thyroid procedures, better outcomes, and lower costs.[50-54] Once complete evaluation of a thyroid nodule suggests a need for surgery, management by a high-volume thyroid cancer surgeon is associated with lower postoperative complication rates.[52]

Perioperative management (Figure 2)

The surgical team must educate the patient about the surgical risks of thyroidectomy, including voice alteration and parathyroid dysfunction. The team should also formulate a comprehensive care plan with anesthesia and perioperative nursing care teams.

Patients should undergo a total or near total thyroidectomy, unless contraindicated, if there is a preoperative diagnosis of differentiated thyroid cancer and any of the following features: lesion > 4 cm, aggressive histologic subtypes, clear extrathyroidal extension, or preoperatively identified nodal metastases. The surgeon may offer less than total thyroidectomy for definitive management to patients with cancers that are between 1 and 4 cm in size without obvious extrathyroidal extension or preoperatively identified nodal metastasis. The surgeon may consider total thyroidectomy in patients where there are contralateral nodules or in patients where postoperative radioactive iodine ablative therapy and/or whole-body scanning and thyroglobulin-based follow-up protocols may be beneficial after a discussion of risks, benefits, and other clinical considerations.

Therapeutic central compartment (level VI and VII) nodal dissection should be carried out in patients with either pre- or intraoperative evidence of gross or radiographically diagnosed lymphadenopathy in these regions. Prophylactic central compartment neck dissection should be considered for patients with T3 or T4 cancers or in the presence of lateral neck metastasis (N1b). Bilateral prophylactic central neck dissection in the absence of either pre- or intraoperative lymphadenopathy is generally not recommended due to a higher risk of post-operative complications that include hypoparathyroidism and RLN injury.

Thyroid lobectomy without central compartment neck dissection should be considered in (cN0) patients with noninvasive T1 or T2 cancers or minimally invasive follicular carcinoma. Any intraoperative findings that suggest more advanced disease should guide surgical decision-making.

A selective therapeutic lateral neck dissection should be performed for patients with DTC and clinically apparent N1b disease. Levels II through V should be considered for dissection, although removal of level Va may be omitted in the absence of clinically or radiographically apparent nodes at this level. When nodes are clinically or radiographically abnormal, they should be included within the field of dissection even if located outside these regions.

The surgeon should offer completion thyroidectomy to patients that have undergone a lobectomy when final pathologic features would have mandated total thyroidectomy had they been known preoperatively. Dissection of central compartment lymph nodes on the previously operated side is generally not necessary without clinical or radiographic evidence of disease. Radioactive iodine (RAI) remnant ablation can be considered in uncommon circumstances when completion thyroidectomy is contraindicated or if the patient refuses further surgery.

Without exception, the recurrent (or non-recurrent) laryngeal nerve should be identified and preserved during thyroidectomy. Care should also be taken to preserve the external branch of the superior laryngeal nerve during thyroidectomy. Nerve monitoring during thyroid operations has utility in patients with invasive or bulky primary or nodal disease, with preoperative vocal fold dysfunction, undergoing revision surgery or in patients requiring bilateral thyroid surgery. In the latter instance, neural monitoring can aid in nerve identification as well as provide prognostic information regarding neural function and is supported by the ATA, AHNS, and American Academy of Otolaryngology Head and Neck Surgery guidelines.

Care should be taken to preserve the parathyroid glands and their blood supply during thyroidectomy. When the glands are removed or separated from their blood supply, confirmed cancer-free grossly normal parathyroid tissue should be autotransplanted into nearby muscle.

Statement 7. To have an understanding of true postoperative voice status following thyroid surgery, a preoperative assessment of the voice and possible laryngeal exam if there is a voice issue or previous neck or chest surgery is recommended. Reported prevalence rates for voice and swallowing symptoms in preoperative thyroidectomy patients range from 11-88%.[55] Although laryngopharyngeal reflux may account for many of these symptoms,[55] all patients scheduled for thyroid surgery should undergo preoperative screening for vocal fold dysfunction by voice or laryngeal examination. Laryngoscopy is mandatory for patients with abnormal findings on screening vocal examination, a history of previous neck or mediastinal surgery, or bulky primary thyroid cancers or nodal disease.

Statement 8. Antibiotics are not usually recommended for clean surgeries including thyroidectomy. Monitoring for antibiotic usage is a way to evaluate quality of care and appropriate utilization. Systemic administration of prophylactic antimicrobials has not been proven effective in reducing surgical site infection rates in patients undergoing clean procedures of the head and neck, including thyroidectomies, and are not recommended for routine use.[56]

Postoperative management (Figure 3)

Following thyroidectomy, important intraoperative findings and details of postoperative care should be shared with the patient and members of the multidisciplinary care team. These include: 1) anatomic findings during surgery, including RLN and parathyroid status, 2) extent of primary and regional nodal disease including the presence of gross residual disease and the presence of gross extrathyroidal extension found at surgery, 3) postoperative status of the voice and laryngeal exam, 4) postoperative serum calcium and parathyroid hormone (PTH) levels and utilization of calcium and/or vitamin D supplementation, and 5) the surgical pathology report.

Statement 9. Monitoring postoperative serum thyroglobulin levels for patients on thyroid hormone therapy or after TSH stimulation is helpful in assessing the persistence of disease or thyroid remnant and predicting future disease recurrence. The postoperative assessment by the endocrinologist or other health care professional is important to measure postoperative thyroid hormone levels and replace thyroid hormone as indicated and to measure postoperative thyroglobulin (Tg) and Tg antibody levels. The latter, along with surgical pathology information, may guide decisions regarding postoperative RAI ablation/treatment and additional imaging. Following lobectomy, a decision may be made in some cases to proceed with completion thyroidectomy to facilitate systemic treatment with RAI while surveillance alone after lobectomy may be appropriate for certain low risk cancers in other cases.

Statement 10. The AJCC Staging and Dynamic Risk Stratification systems for DTC bring value when predicting disease mortality or recurrence, as well as for guiding decisions about treatment and surveillance. A substantial proportion of patients with DTC who are initially classified as intermediate and high risk have an excellent response to therapy and become low risk for developing recurrent disease.[57] Initial risk estimates should be continually modified during ongoing follow up in a dynamic process of risk stratification. The risk of cancer recurrence and disease-specific mortality may change over time as a function of a specific patients' clinical course and response to therapy.

Statement 11. Patients should have their vocal fold function evaluated between 2 weeks and 2 months after thyroid surgery. Among nearly 28,000 patients undergoing surgery for DTC or medullary thyroid cancer, postoperative vocal cord paralysis occurred in 4.1% of instances.[58] Voice assessment should be based on the patient's subjective report and the physician's objective findings. The panel selected this time frame to avoid early assessment which might give a false positive and a hard stop at two months because that should be sufficient time for recovery and evaluation. An abnormal voice evaluation should lead to a formal laryngeal examination to assess vocal fold mobility. Early detection of vocal fold motion abnormalities after thyroidectomy facilitates prompt intervention through vocal fold medialization, improving patient quality of life and diminishing aspiration risk.

Statement 12. For patients who require high dose calcium and vitamin D supplementation at 12 months or longer following total or completion thyroidectomy, a calcium and/or PTH level should be checked. Hypoparathyroidism is the most common complication of total thyroidectomy. Transient hypoparathyroidism occurs in 10% of patients,[59] and long-term outcomes for permanent hypoparathyroidism are measured by obtaining a PTH level longer than 12 months after surgery. Consistent measurement enables quality improvement.

Statement 13. Mortality rates should be monitored after thyroid surgery.

Statement 14. Readmission rates within the first 30 days following thyroid surgery is a potential proxy for some complications after surgery.

Statement 15. Reoperation rate during the first 30 days is an indicator for some postoperative surgical complications after thyroid surgery (e.g., hematoma, vocal cord paralysis and aspiration requiring medialization).

Statement 16. The average length of stay is a good proxy for perioperative resource management and allows comparisons to be made to other surgical and medical patients.

In addition to RLN and parathyroid postoperative assessment, other more global measures of perioperative quality of care should be assessed. Multiple studies examining quality and safety in thyroid surgery assess mortality,[53,60-64] readmissions with and without related data on reoperations,[60,62,63,65-68] and length of stay (LOS).[50-53,64,69] Recommended quality metrics include mortality rates, readmission rates, reoperation rates, and LOS in hospital.

After considering comorbidities, available social support, perioperative education, and proximity to emergency services, same-day surgery may be a reasonable alternative to overnight observation. Same-day surgery may not be possible in all clinical settings; consequently, the consensus group selected LOS as a quality measure.

DISCUSSION

Through broad multispecialty consensus representation from societies in the U.S. and Canada, we developed both workflows and quality statements for thyroid nodule care. While clinical practice guidelines (CPG) articulate current best evidence and are designed to reduce unwarranted variation in care that can negatively affect quality outcomes, they often benefit from efforts to support dissemination and implementation. The workflows and statements presented here were created for just that purpose and are intended for use across the range of practice settings in North America.

Implementing workflows and metrics

The integration of practice algorithms and quality metrics such as those outlined here is a complex process. It requires extensive collaboration between primary care providers, endocrinologists, thyroid surgeons, radiologists, nuclear medicine physicians, radiation oncologists and pathologists. At KPNC and CCO, there was strong executive and leadership support for implementation. At both organizations, it was a group process: multidisciplinary teams and clinical champions worked together to review the evidence over several months meeting bimonthly face-to-face for a full day to draft the work flows and determine responsibilities for the various processes of care. Compromises were required to respect the perspectives of all involved specialties and to accommodate variations, such as whether an endocrinologist, surgeon or radiologist performed ultrasound for risk stratification and assessment of the thyroid, central, and lateral neck. Clinical champions then took responsibility for helping with dissemination, addressing concerns of end users in their specialty, urging their acceptance, and keeping them in active use.

Education was achieved through multiple in-person or webinar CME sessions to educate providers and obtain feedback. Robust discussion occurred at these events, allowing resistance to change to be expressed, acknowledged, and addressed. The feedback was used to further refine the workflows during the rollout. As implementation continued, additional CME events were modified to reflect any changes in the workflows, reinforce collaboration across the specialties, and demonstrate how workflows are based on clinical practice guidelines. At KPNC, an in-house medical journal published the workflows to facilitate their distribution prior to the launch.[23] In addition, after implementation at KPNC, a performance improvement team using tools based on human-centered design interviewed physicians, nurses, and patients to further assure a collaborative approach to the workflows and patient-centered decision making regarding disposition following surgery, including same-day discharge. CCO completed review in its 14 regions before finalizing and distributing care paths though Community of Practice events.

Using quality metrics

This consensus statement includes quality metric statements because setting and reviewing local goals is key for driving ongoing communication and improving quality of patient care. The selected quality metric statements focus on key concepts, and we explicitly did not define numerators, denominators, or equations for each metric because we recognize that individual sites should be able to adapt their use to local constraints. Once integrated into practice, institutions may wish to monitor the success of the implementation on a quarterly basis, providing feedback to individual surgeons and measuring their institution against others. The metrics can identify both superior performance as well as opportunities for improvement. Institutions can also use the data to inform patients about case volumes and outcomes.

CONCLUSIONS

With these workflows and quality metric statements the American Head and Neck Society Endocrine Section provides a simplified approach to incorporating the 2015 ATA guidelines into the everyday management of thyroid nodules and DTC and have the potential to improve quality and decrease unwarranted variations in care. For implementation, we suggest that users create multidisciplinary teams in their local settings to review, refine, implement, and sustain these practices.

Disclaimer

The American Thyroid Association (ATA) develops Clinical Practice Guidelines to provide guidance and recommendations for practice areas concerning thyroid disease and thyroid cancer. The Guidelines are not inclusive of all proper approaches or methods, or exclusive of others. The Guidelines do not establish a standard of care and specific outcomes are not guaranteed.

A guideline is not intended to take the place of physician judgment in diagnosing and treatment of patients. It is also not intended to serve as a basis to approve or deny financial coverage for any specific therapeutic or diagnostic modality. Treatment decisions must be made based on the independent judgment of health care providers and each patient's individual circumstances.

The ATA develops guidelines based on the evidence available in the literature and the expert opinion of the task force in the recent timeframe of the publication of the guidelines. Management issues have not been and cannot be comprehensively addressed in randomized trials; therefore, the evidence cannot be comprehensive. Guidelines cannot always account for individual variation among patients. Guidelines cannot be considered inclusive of all proper methods of care or exclusive of other treatments reasonably directed at obtaining the same results.

Therefore, the ATA considers adherence to the its guideline to be voluntary, with the ultimate determination regarding its application to be made by the treating physician and health care professionals with the full consideration of the individual patient's clinical history and physical status. In addition, the guideline concerns the therapeutic interventions used in clinical practice and do not pertain to clinical trials. Clinical trials are a separate matter, designed to research new and novel therapies, and the guidelines are not necessarily relevant to their purpose.

Guideline development includes an identification of areas for future study and research, indicating the focus for future investigational therapy; based on the findings reviewed and synthesized from the latest literature.

Acknowledgements

Angelo Santoni, Sushma Shankar, John L. Adams of Kaiser Permanente provided data analysis. Jenni Green of Kaiser Permanente provided writing and editing support.

REFERENCES

1. Vander JB, Gaston EA, Dawber TR. The significance of nontoxic thyroid nodules. Final report of a 15-year study of the incidence of thyroid malignancy. Ann Intern Med 1968;69:537-40.

2. Tunbridge WM, Evered DC, Hall R, et al. The spectrum of thyroid disease in a community: the Whickham survey. Clin Endocrinol (Oxf) 1977;7:481-93.

3. Tamhane S, Gharib H. Thyroid nodule update on diagnosis and management. Clin Diabetes Endocrinol 2016;2:17.

4. Guth S, Theune U, Aberle J, Galach A, Bamberger CM. Very high prevalence of thyroid nodules detected by high frequency (13 MHz) ultrasound examination. Eur J Clin Invest 2009;39:699-706.

5. Haugen BR, Alexander EK, Bible KC, et al. 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. Thyroid 2016;26:1-133.

6. Hegedus L. Clinical practice. The thyroid nodule. N Engl J Med 2004;351:1764-71.

7. Iglesias ML, Schmidt A, Ghuzlan AA, et al. Radiation exposure and thyroid cancer: a review. Arch Endocrinol Metab 2017;61:180-7.

8. Werk EE, Jr., Vernon BM, Gonzalez JJ, Ungaro PC, McCoy RC. Cancer in thyroid nodules. A community hospital survey. Arch Intern Med 1984;144:474-6.

9. Lin JD, Chao TC, Huang BY, Chen ST, Chang HY, Hsueh C. Thyroid cancer in the thyroid nodules evaluated by ultrasonography and fine-needle aspiration cytology. Thyroid 2005;15:708-17.

10. Xu L, Li G, Wei Q, El-Naggar AK, Sturgis EM. Family history of cancer and risk of sporadic differentiated thyroid carcinoma. Cancer 2012;118:1228-35.

11. Sherman SI. Thyroid carcinoma. Lancet 2003;361:501-11.

12. Aschebrook-Kilfoy B, Schechter RB, Shih YC, et al. The clinical and economic burden of a sustained increase in thyroid cancer incidence. Cancer Epidemiol Biomarkers Prev 2013;22:1252-9.

13. Lim H, Devesa SS, Sosa JA, Check D, Kitahara CM. Trends in thyroid cancer incidence and mortality in the United States, 1974-2013. JAMA 2017;317:1338-48.

14. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2016. CA Cancer J Clin 2016;66:7-30.

15. Brito JP, Al Nofal A, Montori VM, Hay ID, Morris JC. The impact of subclinical disease and mechanism of detection on the rise in thyroid cancer incidence: a population-based study in Olmsted County, Minnesota during 1935 through 2012. Thyroid 2015;25:999-1007.

16. Davies L, Welch HG. Increasing incidence of thyroid cancer in the United States, 1973-2002. JAMA 2006;295:2164-7.

17. Morris LG, Sikora AG, Tosteson TD, Davies L. The increasing incidence of thyroid cancer: the influence of access to care. Thyroid 2013;23:885-91.

18. Francis DO, Randolph G, Davies L. Nationwide variation in rates of thyroidectomy among US Medicare beneficiaries. JAMA Otolaryngol Head Neck Surg 2017;143:1122-5.

19. Field MJ, Lohr KN, eds. Guidelines for clinical practice: from development to use. Washington, DC: Institute of Medicine, National Academies Press; 1992.

20. Cabana MD, Rand CS, Powe NR, et al. Why don't physicians follow clinical practice guidelines? A framework for improvement. JAMA 1999;282:1458-65.

21. Tu SW, Musen MA. A typology for modeling processes in clinical guidelines and protocols. Communications in Computer and Information Science. Heidelberg: Springer Berlin Heidelberg; 2010. p. 545.

22. A coordinated approach to improve consistency in surgical care across Canada: pan-Canadian standards for eight types of cancer. Canadian Partnership Against Cancer, 2016. (Accessed April 23, 2018, at https://content.cancerview.ca/download/cv/treatment_and_support/diagnosis_and_treatment/documents/pan_canadian_standards_eight_types_cancer_reportpdf?attachment=0.)

23. Meltzer C, Budayr A, Chavez A, et al. Evidence-based workflows for thyroid and parathyroid surgery. Perm J 2016;20:57-73.

24. Qaseem A, Snow V, Owens DK, Shekelle P, Clinical Guidelines Committee of the American College of Physicians. The development of clinical practice guidelines and guidance statements of the American College of Physicians: summary of methods. Ann Intern Med 2010;153:194-9.

25. Schunemann HJ, Oxman AD, Brozek J, et al. Grading quality of evidence and strength of recommendations for diagnostic tests and strategies. BMJ 2008;336:1106-10.

26. Brozek JL, Akl EA, Jaeschke R, et al. Grading quality of evidence and strength of recommendations in clinical practice guidelines: part 2 of 3. The GRADE approach to grading quality of evidence about diagnostic tests and strategies. Allergy 2009;64:1109-16.

27. Bossuyt PM, Reitsma JB, Standards for Reporting of Diagnostic Accuracy Group. The STARD initiative. Lancet 2003;361:71.

28. Leung AN, Bull TM, Jaeschke R, et al. An official American Thoracic Society/Society of Thoracic Radiology clinical practice guideline: evaluation of suspected pulmonary embolism in pregnancy. Am J Respir Crit Care Med 2011;184:1200-8.

29. Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med 2011;155:529-36.

30. Stack BC, Jr., Bimston DN, Bodenner DL, et al. American Association of Clinical Endocrinologists and American College of Endocrinology disease state clinical review: postoperative hypoparathyroidism--definitions and management. Endocr Pract 2015;21:674-85.

31. Mechanick JI, Camacho PM, Garber AJ, et al. American Association of Clinical Endocrinologists and American College of Endocrinology protocol for standardized production of clinical practice guidelines, algorithms, and checklists--2014 update and the AACE G4G Program. Endocr Pract 2014;20:692-702.

32. Mechanick JI, Pessah-Pollack R, Camacho P, et al. American Association of Clinical Endocrinologists and American College of Endocrinology protocol for standardized production of clinical practice guidelines, algorithms, and checklists - 2017 update. Endocr Pract 2017;23:1006-21.

33. Mechanick JI, Bergman DA, Braithwaite SS, Palumbo PJ, American Association of Clinical Endocrinologists Ad Hoc Task Force for Standardized Production of Clinical Practice Guidelines. American Association of Clinical Endocrinologists protocol for standardized production of clinical practice guidelines. Endocr Pract 2004;10:353-61.

34. Mechanick JI, Camacho PM, Cobin RH, et al. American Association of Clinical Endocrinologists protocol for standardized production of clinical practice guidelines--2010 update. Endocr Pract 2010;16:270-83.

35. The RAND/UCLA appropriateness method user's manual. Rand Corporation, 2001. (Accessed April 23, 2018, at https://www.rand/org/pubs/monograph_reports/MR1269.html.)

36. Pincus T, Miles C, Froud R, Underwood M, Carnes D, Taylor SJ. Methodological criteria for the assessment of moderators in systematic reviews of randomised controlled trials: a consensus study. BMC Med Res Methodol 2011;11:14.

37. Dalkey N, Helmer O. An experimental application of the Delphi method in the use of experts. Manage Sci 1963;9:458-67.

38. The Delphi method: an experimental study of group opinion. RAND Corporation, 1969. (Accessed April 23, 2018, at https://www.rand.org/pubs/research_memoranda/RM5888.html.)

39. Basger BJ, Chen TF, Moles RJ. Validation of prescribing appropriateness criteria for older Australians using the RAND/UCLA appropriateness method. BMJ Open 2012;2.

40. Pre-analysis plan: decision rules for adolescent substance use treatment settings: an online Delphi process. RAND Corporation, 2017. (Accessed April 23, 2018, at https://osf.io/c5ffh/?action=download&version=1.)

41. Tessler FN, Middleton WD, Grant EG, et al. ACR Thyroid Imaging, Reporting and Data System (TI-RADS): white paper of the ACR TI-RADS Committee. J Am Coll Radiol 2017;14:587-95.

42. Martin D, Miller AP, Quesnel-Vallee A, Caron NR, Vissandjee B, Marchildon GP. Canada's universal health-care system: achieving its potential. Lancet 2018;391:1718-35.

43. Eskander A, Devins GM, Freeman J, et al. Waiting for thyroid surgery: a study of psychological morbidity and determinants of health associated with long wait times for thyroid surgery. Laryngoscope 2013;123:541-7.

44. Gharib H, Papini E, Garber JR, et al. American Association of Clinical Endocrinologists, American College of Endocrinology, and Associazione Medici Endocrinologi medical guidelines for clinical practice for the diagnosis and management of thyroid nodules--2016 update. Endocr Pract 2016;22:622-39.

45. Davies L, Welch HG. Current thyroid cancer trends in the United States. JAMA Otolaryngol Head Neck Surg 2014;140:317-22.

46. Cibas ES, Ali SZ. The 2017 Bethesda System for Reporting Thyroid Cytopathology. Thyroid 2017;27:1341-6.

47. Lee S, Skelton TS, Zheng F, et al. The biopsy-proven benign thyroid nodule: is long-term follow-up necessary? J Am Coll Surg 2013;217:81-8; discussion 8-9.

48. Mazotas IG, Wang TS. The role and timing of parathyroid hormone determination after total thyroidectomy. Gland Surg 2017;6:S38-S48.

49. Helme S, Lulsegged A, Sinha P. Incidental parathyroid disease during thyroid surgery: should we remove them? ISRN Surg 2011;2011:962186.

50. Al-Qurayshi Z, Robins R, Hauch A, Randolph GW, Kandil E. Association of surgeon volume with outcomes and cost savings following thyroidectomy: a national forecast. JAMA Otolaryngol Head Neck Surg 2016;142:32-9.

51. Hauch A, Al-Qurayshi Z, Randolph G, Kandil E. Total thyroidectomy is associated with increased risk of complications for low- and high-volume surgeons. Ann Surg Oncol 2014;21:3844-52.

52. Sosa JA, Bowman HM, Tielsch JM, Powe NR, Gordon TA, Udelsman R. The importance of surgeon experience for clinical and economic outcomes from thyroidectomy. Ann Surg 1998;228:320-30.

53. Sosa JA, Mehta PJ, Wang TS, Boudourakis L, Roman SA. A population-based study of outcomes from thyroidectomy in aging Americans: at what cost? J Am Coll Surg 2008;206:1097-105.

54. Adam MA, Thomas S, Youngwirth L, et al. Is there a minimum number of thyroidectomies a surgeon should perform to optimize patient outcomes? Ann Surg 2017;265:402-7.

55. Holler T, Anderson J. Prevalence of voice & swallowing complaints in pre-operative thyroidectomy patients: a prospective cohort study. J Otolaryngol Head Neck Surg 2014;43:28.

56. Bratzler DW, Dellinger EP, Olsen KM, et al. Clinical practice guidelines for antimicrobial prophylaxis in surgery. Am J Health Syst Pharm 2013;70:195-283.

57. Krajewska J, Chmielik E, Jarzab B. Dynamic risk stratification in the follow-up of thyroid cancer: what is still to be discovered in 2017? Endocr Relat Cancer 2017;24:R387-R402.

58. Papaleontiou M, Hughes DT, Guo C, Banerjee M, Haymart MR. Population-based assessment of complications following surgery for thyroid cancer. J Clin Endocrinol Metab 2017;102:2543-51.

59. Khan MI, Waguespack SG, Hu MI. Medical management of postsurgical hypoparathyroidism. Endocr Pract 2011;17 Suppl 1:18-25.

60. Hessman C, Fields J, Schuman E. Outpatient thyroidectomy: is it a safe and reasonable option? Am J Surg 2011;201:565-8.

61. Kandil E, Noureldine SI, Abbas A, Tufano RP. The impact of surgical volume on patient outcomes following thyroid surgery. Surgery 2013;154:1346-52; discussion 52-3.

62. Khavanin N, Mlodinow A, Kim JY, Ver Halen JP, Antony AK, Samant S. Assessing safety and outcomes in outpatient versus inpatient thyroidectomy using the NSQIP: a propensity score matched analysis of 16,370 patients. Ann Surg Oncol 2015;22:429-36.

63. Orosco RK, Lin HW, Bhattacharyya N. Ambulatory thyroidectomy: a multistate study of revisits and complications. Otolaryngol Head Neck Surg 2015;152:1017-23.

64. Stavrakis AI, Ituarte PH, Ko CY, Yeh MW. Surgeon volume as a predictor of outcomes in inpatient and outpatient endocrine surgery. Surgery 2007;142:887-99; discussion -99.

65. Houlton JJ, Pechter W, Steward DL. PACU PTH facilitates safe outpatient total thyroidectomy. Otolaryngol Head Neck Surg 2011;144:43-7.

66. Khavanin N, Mlodinow A, Kim JY, Ver Halen JP, Samant S. Predictors of 30-day readmission after outpatient thyroidectomy: an analysis of the 2011 NSQIP data set. Am J Otolaryngol 2014;35:332-9.

67. Mullen MG, LaPar DJ, Daniel SK, Turrentine FE, Hanks JB, Smith PW. Risk factors for 30-day hospital readmission after thyroidectomy and parathyroidectomy in the United States: an analysis of National Surgical Quality Improvement Program outcomes. Surgery 2014;156:1423-30; discussion 30-1.

68. Terris DJ, Snyder S, Carneiro-Pla D, et al. American Thyroid Association statement on outpatient thyroidectomy. Thyroid 2013;23:1193-202.

69. Hauch A, Al-Qurayshi Z, Friedlander P, Kandil E. Association of socioeconomic status, race, and ethnicity with outcomes of patients undergoing thyroid surgery. JAMA Otolaryngol Head Neck Surg 2014;140:1173-83.