Cost-Effectiveness of the ACR TIRADS Compared to the ATA 2015 Risk Stratification Systems in the Evaluation of Incidental Thyroid Nodules

Emily HT Pang, Alexander W. Carter, Eitan Prisman, Stuart H. Kreisman

Rationale and Objectives: Thyroid nodules are a common incidental imaging finding and prone to overdiagnosis. Several risk stratification systems have been developed to reduce unnecessary work-up, with two of the most utilized including the American Thyroid Association 2015 (ATA2015) and the newer American College of Radiology Thyroid Imaging, Reporting and Data System (TIRADS) guidelines. The purpose of this study is to evaluate the cost-effectiveness of the ATA2015 versus the TIRADS guidelines in the management of incidental thyroid nodules.

Methods: A cost-utility analysis was conducted using decision tree modeling, evaluating adult patients with incidental thyroid nodules < 4 cm. Model inputs were populated using published literature, observational data, and expert opinion. Single-payer perspective, Canadian dollar currency, five-year time horizon, willingness to pay (WTP) threshold of $50,000, and discount rate of 1.5% per annum were utilized. Scenario, deterministic and probabilistic sensitivity analyses were performed. The primary outcome was the incremental cost-effectiveness ratio (ICER) expressed as incremental cost per quality-adjusted life year (QALY) gained.

Results: For the base case scenario, TIRADS dominated the ATA2015 strategy by a slim margin, producing 0.005 more QALYs at $25 less cost. Results were sensitive to the malignancy rate of biopsy and the utilities of a patient with a benign nodule/subclinical malignancy or under surveillance. Probabilistic sensitivity analysis showed that TIRADS was the more cost-effective option 79.7% of the time.

Conclusion: The TIRADS guidelines may be the more cost-effective strategy by a small margin compared to ATA2015 in most scenarios when used to risk stratify incidental thyroid nodules.

Key Words: Thyroid nodule; TIRADS; Ultrasound; Cost-effectiveness.

INTRODUCTION

Thyroid nodules are commonly detected as incidental findings on imaging studies performed for other reasons, perhaps unsurprisingly as the estimated prevalence of thyroid nodules in the general population is up to 65% (1). While the risk of malignancy is low (approximately 2–12%) (2–5) upon detection, radiologists are generally obligated to suggest dedicated work-up in many cases, usually with ultrasound. Unfortunately, it can be challenging to differentiate between benign and malignant nodules with certainty on ultrasound alone, and definitive diagnosis requires a fine needle aspiration (FNA) biopsy procedure. Most FNA biopsies will end up yielding benign cytology (6) and even when a thyroid malignancy is detected, in the majority of cases it follows an indolent course (7) and will not result in significant morbidity or mortality in a patient's lifetime. Over the last 20–30 years, an increase in incidental thyroid nodule and thyroid cancer detection has paralleled a general increase in medical imaging utilization, contributing to what has been deemed an "epidemic of overdiagnosis of thyroid cancer"—an increase in incidence without any substantial increase in mortality (8). Concerns regarding overdiagnosis relate to consumption of healthcare resources such as tests and procedures (for example, thyroid FNA or surgery) which do not ultimately improve a patient's health.

Acad Radiol 2024; 31:3993–4010

From the Department of Radiology, Vancouver General Hospital, University of British Columbia, Vancouver, BC, Canada (E.H.P.); Department of Health Policy, London School of Economics, London, England, UK (A.W.C.); Division of Otolaryngology Head and Neck Surgery, Department of Surgery, Vancouver General Hospital, University of British Columbia, Vancouver, BC, Canada (E.P.); Division of Endocrinology, Department of Internal Medicine, St Paul's Hospital, University of British Columbia, Vancouver, BC, Canada (S.H.K.). Received December 11, 2023; revised April 5, 2024; accepted April 5, 2024. **Address correspondence to**: E.H.T.P. e-mail: emily.pang@vch.ca

https://doi.org/10.1016/j.acra.2024.04.004

The workup of thyroid nodules consumes a not insignificant proportion of ultrasound department resource usage, with diagnostic thyroid ultrasounds comprising approximately 5–10% of daily volume at our institution and each FNA biopsy requiring at least 30 min of radiologist and assistant time. As most of these biopsies will ultimately be negative for malignancy, this raises the question of how we can better risk-stratify these thyroid nodules to reduce unnecessary FNA. There have been several risk stratification system (RSS) guidelines developed which have attempted to address this problem. Up until recently, the 2015 American Thyroid Association (ATA) guidelines (ATA2015) were widely used, proposing a "level of suspicion" hierarchy based on the nodule pattern on ultrasound with specific size thresholds for FNA at each level (9). The ATA2015 guidelines can be unwieldy for practical use, however, and offered vague guidance on how to handle nodules which did not meet biopsy thresholds.

In 2017, the American College of Radiology published the Thyroid Imaging Reporting and Data System (TIRADS) guidelines, which utilizes a point-based system to place a nodule into a TIRADS category from 1 to 5 based on ultrasound features, and provides specific guidance as to when to perform FNA (generally narrower criteria compared to ATA) as well as when and how to manage surveillance of nodules which did not meet FNA criteria (10). A brief overview of the two guidelines is provided in Table 1. Due to its ease of use, TIRADS guidelines have been increasingly adopted in many North American radiology practices (11) and in many cases supplanted the ATA guidelines. However, it is not without its criticisms—in particular the concern that the trade-off of fewer biopsies would be a costly explosion of ultrasound surveillance for little benefit as well as greater potential for missed or delayed malignancy diagnoses (12). However, it is also possible that TIRADS could ultimately conserve resources by reducing both biopsies and subsequent investigations/surgeries of questionable necessity for a low-risk malignancy (13).

As the TIRADS system is still relatively new, there is not yet robust evidence to support or refute these criticisms; its potential economic impact on the healthcare system—especially compared to the ATA2015 guidelines—is still an open question. To our knowledge to date, there have been no published economic evaluations which specifically address this issue. The purpose of this study is to evaluate the cost-effectiveness of the TIRADS versus ATA2015 guidelines in the management of incidental thyroid nodules in adult patients in the Canadian setting from a provincial (British Columbia, BC) public single-payer perspective. We hypothesize that TIRADS is the more cost-effective option; the costs of increased surveillance ultrasounds prescribed by TIRADS are outweighed by the reduction in costs of potentially unnecessary procedures without significant decrement in quality of life.

METHODS

Study Design

Institutional ethics approval was obtained for this study. This study is a decision tree model-based cost-utility analysis, comparing the TIRADS versus the ATA2015 strategies to the primary outcome being the incremental cost per quality-adjusted-life-year (QALY) gained (the denominator QALY used to assess the patient health impacts). The QALY is a measure of disease impact which includes the length of life adjusted by a quality of life modifier (utility), with 1 equating to a year of perfect health and 0 equating to death by convention. A decision tree model structure (Fig 1) was chosen to reflect the clinical algorithm for both strategies (see Appendix) and was developed in consultation with stakeholder clinicians in the fields of endocrine surgery and endocrinology. A do-nothing alternative was not considered due to potential ethical questions and limited natural history literature. A publicly funded single payer perspective (Canadian province of British Columbia Ministry of Health) is utilized for analysis, with inclusion of direct medical costs. Future events were discounted at a rate of 1.5% per annum as per Canadian Agency for Drugs and Technologies in Health (CADTH) recommendations (14). Standard software (Treeage Pro 2022) was used to construct the model and was also used to perform the base-case, scenario, and sensitivity analyses.

The base case population is asymptomatic adult patients (18 yo and older) undergoing ultrasound assessment in British Columbia (BC), Canada for an incidentally discovered thyroid nodule measuring less than 4 cm on imaging performed for other reasons. The nodule has no definitive aggressive features such as extrathyroidal extension or neck adenopathy. There are no thyroid cancer risk factors such as a previous history of thyroid malignancy or neck radiation, and there are no comorbidities which would preclude surgery. The time horizon for the model is 5 years as this is the maximum surveillance length for uncomplicated nodules under TIRADS and while ATA2015 does not prescribe a surveillance length, it does note that there are "no follow-up studies of nodule growth that extend observation beyond 5 years" (9) to inform the guidelines.

Model Assumptions

For the base case population, any patients with increased thyroid malignancy risk from baseline (i.e. prior history of thyroid cancer, neck radiation, FDG-avid nodules, neck adenopathy) or other factors which may complicate the treatment decision such as the presence of symptoms or comorbidities precluding surgery were excluded as the RSS guidelines may not apply. Nodules greater than 4 cm were also excluded as RSS guidelines may be irrelevant since surgery may be indicated regardless of biopsy result (15). For

TABLE 1. Summary of ATA2015 vs ACR TIRADS Guidelines

ATA2015

Ultrasound Pattern	Recommendation
High suspicion Solid hypoechoic nodule or solid hypoechoic component of a partially cystic nodule with one or more of the following features: irregular margins, microcalcifications, taller than wide shape, rim calcifications with small extrusive soft tissue component, evidence of extrathyroidal extension	FNA at ≥ 1 cm
Intermediate suspicion Hypoechoic solid nodule with smooth margins without high suspicion features.	FNA at ≥ 1 cm
Low suspicion Isoechoic or hyperechoic solid nodule, or partially cystic nodule with eccentric solid areas, without high suspicion features.	FNA at ≥ 1.5 cm
Very low suspicion Spongiform or partly cystic nodules without any of the low, intermediate or high suspicion features.	Consider FNA at ≥ 2 cm *Observation without FNA is also a reasonable option*
Benign Purely cystic nodules	No FNA

ACR TIRADS

Composition (choose 1)	Echogenicity (choose 1)	Shape (choose 1)	Margin (choose 1)	Echogenic Foci (choose all that apply)
– Cystic (0 pt) – Spongiform (0 pt) – Mixed cystic/solid (1 pt) – - Solid (2 pt)	– Anechoic (0 pt) – Hyper/isoechoic (1 pt) – Hypoechoic (2 pt) – Very hypoechoic (3 pt)	– Wider than tall (0 pt) – Taller than wide (3 pt)	– Smooth (0 pt) – Ill-defined (0 pt) – Lobulated or irregular (2 pt) – Extra-thyroidal extension (3 pt)	– None or large comet tail (0 pt) – Macrocalcifications (1 pt) – Rim calcification (2 pt) – Punctate (3 pt)

ADD POINTS

TIRADS 1	*TIRADS 2*	*TIRADS 3*	*TIRADS 4*	*TIRADS 5*
0 points No FNA or surveillance	2 points No FNA or surveillance	3 points FNA at ≥ 1.5 cm Surveillance at ≥ 1 cm at 1,2,3,5 years.	4-6 points FNA at ≥ 1.5 cm Surveillance at ≥ 1 cm at 1,2,3,5 years.	≥ 7 points FNA at ≥ 1 cm Surveillance at ≥ 0.5 cm annually for 5 years

Adapted from (9,10)

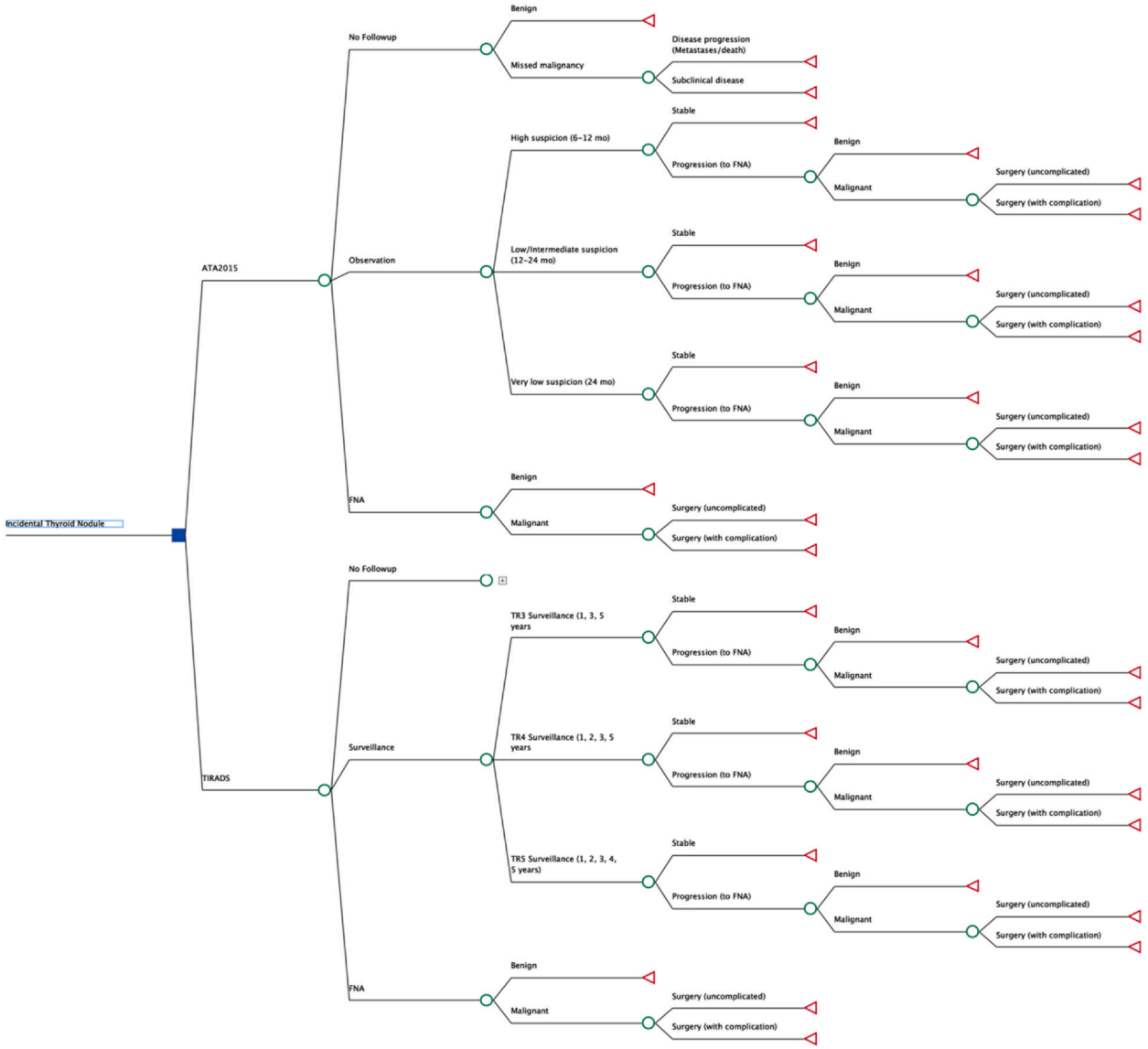


Figure 1. Condensed decision tree model for TIRADS vs ATA2015 risk stratification systems in the assessment of incidental thyroid nodules. ATA, American Thyroid Association; FNA, fine needle aspiration; TIRADS, thyroid imaging reporting and data system. A more detailed version of the model is included in the Appendix.

patients with multiple nodules, only the highest suspicion nodule is considered as this would likely have the most impact on the associated costs and utilities. Additionally, most ATA2015 versus TIRADS head-to-head studies evaluated performance at the individual nodule rather than patient level. Although interobserver variability is a recognized issue (for example, in assigning echogenic foci and echogenicity scores with TIRADS) (16,17), it is also assumed that there is negligible inter- and intra-observer variability in the assignment of TIRADS or ATA2015 classifiers, as this would otherwise be impractical to model directly. However, at least some of this variability is accounted for in the sensitivity analysis as a range of possible nodule assignment proportions is included. Additionally, there is also inherent redundancy in TIRADS and ATA which allows for a range of point assignments to result in the same TIRADS score (for TIRADS 4 or 5) and different nodule patterns resulting in the same ATA risk category which also helps to mitigate this.

ATA2015 does not specify the interval or duration of surveillance for nodules which do not meet FNA criteria. For the base case, a single 1-year surveillance ultrasound was assumed as this was common practice as per expert opinion *(M.Dahl, personal communication, Sept 2022)*. The ATA guidelines also provide an option to FNA biopsy or observe very low suspicion nodules greater than 2 cm—for the base case, it was assumed that these nodules were biopsied as is usually in case in practice as well as the cited literature.

Due to the lack of data on the natural history and real-world management of nodules that are not initially biopsied under each strategy, it is assumed that (1) all nodules eligible

for surveillance complete the initially prescribed surveillance schedule with the exception of nodules that are upgraded to meet FNA threshold, (2) any events including upgrading of nodules under surveillance, as well as death and metastases secondary to missed malignancy in nodules not followed up on average occur halfway through the 5-year time horizon (i.e. middle of year 3), and (3) the proportion of nodules under surveillance upgraded to FNA is the same regardless of initial nodule suspicion pattern (18).

The handling of nondiagnostic (category I result using the standardized Bethesda system for reporting thyroid cytopathology (19)) and indeterminate (Bethesda III and IV) biopsies as well as when to perform a repeat biopsy when there is suspicion for false negative cytology were felt to be beyond the scope of this study for a few reasons:

- These issues are not addressed in the TIRADS white paper
- Some reference studies omitted nondiagnostic and indeterminate biopsies from their analyses (20)
- The rate of nondiagnostic and indeterminate biopsies is not expected to differ dramatically between TIRADS and ATA2015. For example, in a prospective study of 492 nodules, Koc et al (21) found 16.9% yielded Bethesda III cytology when FNA'ed under TIRADS criteria versus 13.5% for ATA2015.
- There is also variability in actual practice with respect to the management of these nodules which is also unlikely to be dependent on whether a nodule was initially assessed with TIRADS or ATA2015.

As there is no evidence that there would be a substantive difference between the two strategies when it comes to indeterminate or nondiagnostic nodules, they would be unlikely to impact the final comparative results. Thus, a benign result on FNA biopsy was considered a terminal result, and both non-diagnostic biopsies and nodules with indeterminate histology were excluded. Of note, the cumulative incidence of Bethesda I, III, and IV results on FNA has been estimated to be between 14–38% (22,23). The risk of malignancy is assumed to be the same for nodules initially selected for FNA as well as those which are upgraded to FNA on surveillance.

For the base case, it is assumed that a hemithyroidectomy is performed for all patients with a suspicious or malignant result on biopsy (i.e. Bethesda category V or VI result). This scenario was chosen as the vast majority of patients (approximately 90%) will initially undergo hemithyroidectomy at our institution, and in general, this approach has become increasingly adopted over the last several years especially in patients without adverse features such as extrathyroidal extension or lymph node metastases (24) as with the base case population. In practice, some patients may undergo an initial total or subsequent completion thyroidectomy depending on a number of factors including surgeon and patient preference, nodule histology, nodal disease, the presence of extrathyroidal extension or molecular testing results (9). Attempting to account for all possibilities would unnecessarily overcomplicate the model, and data on the specific break-down of patients who would be eligible for total or completion thyroidectomy was lacking in studies which directly compared TIRADS and ATA. More importantly, the surgical decision is generally independent of whether ATA or TIRADS was used to initially risk stratify the nodule. The two major potential long-term complications of hemithyroidectomy were factored into the model, namely unilateral recurrent laryngeal nerve (RLN) injury and hypothyroidism. A wide upper range of costs was included in the sensitivity analyses to partly account for the potential increased costs of total/completion thyroidectomy and its complications. The sequence of events from FNA to surgery and any complications are assumed to occur within a year when determining QALYs. This is in line with the interim results of a clinical trial comparing nodule surveillance versus published by Wong et al which reported that the majority of interventions occurred during the 12-month follow-up period (25,26).

Model Input Parameters

Probabilities

To populate the model probabilities, observational data were collected from our institution and a systematic literature review was performed (see Appendix). The primary motivation for the observational data collection was to supplement the lack of published data on (1) the distribution of ATA2015 and TIRADS management assignments in a population which only includes incidentally detected nodules and (2) the probability of a nodule initially assigned to surveillance by ATA2015 or TIRADS progressing to meet FNA thresholds during the surveillance period. Results were pooled from eligible studies + /- observational data (where appropriate) to determine the base case probabilities.

For the observational component, retrospective data from 200 consecutive patients who underwent initial ultrasound evaluation for an incidental thyroid nodule at our institution (a tertiary care academic hospital) from February 2018 to March 2019 was collected. Only the highest suspicion nodule (highest TIRADS point score, with larger size as tie-breaker if required) was considered for each patient. For each nodule, the following data was recorded:

- The TIRADS score originally assigned
- An ATA2015 classification retrospectively assigned in consensus by two experienced ultrasound radiologists blinded to the initial TIRADS score.
- Whether the nodule met FNA, surveillance, or no follow-up criteria as per TIRADS and ATA2015 guidelines

For nodules that met surveillance criteria and had follow-up ultrasounds available (up to 5 years from the initial ultrasound in line with the time horizon of the study), any that were subsequently upgraded to meet FNA biopsy thresholds

were also recorded. A summary of observational data results is included in the Appendix.

For the systematic literature search, a Pubmed search was performed using the search string "(thyroid) AND ((thyroid imaging reporting and data system) OR (TIRADS) OR (TI-RADS)) AND ((American Thyroid Association) OR (ATA))". Only studies on adult patients which directly compared ATA2015 and TIRADS guidelines for the same population of nodules were included. Studies at high risk for selection bias (for example, non-consecutive patients, only including patients who underwent surgery or referred to an oncology center), which did not have English full-text available, or had insufficiently detailed results to calculate probabilities were excluded. Other probabilities related to surgical outcomes were obtained from a separate thorough literature search.

Costs

All costs are in 2021 Canadian dollars, adjusted for inflation when required using the Bank of Canada Inflation Calculator (27). Costing methodology was performed in accordance with CADTH guidelines (28). Hospital costs were estimated from the Canadian Institute of Health Information (CIHI) patient cost estimator using Case Mix Group (CMG) code 424 (Thyroid/Parathyroid/Thymus Intervention) for BC (29). Physician (including consult, diagnostic imaging, and procedural) costs and outpatient laboratory costs were derived from the British Columbia Ministry of Health published payment schedules (30,31).

The included costs of FNA biopsy in BC include the pre-procedural ultrasound, mini tray fee, physician fee (both for the diagnostic ultrasound component and procedural component), and cytology. The included costs of surgery assumes a typical workup for a patient undergoing hemithyroidectomy at our institution as well as costs associated with the two major long-term surgical complications (i.e. hypothyroidism, unilateral RLN injury) including medications, follow-up appointments and procedures). Indirect medical costs were not included. A detailed accounting of included costs is included in the Appendix.

Utilities

Utility values, ranges for sensitivity analyses, and distributions for health state utilities were sourced from a thorough review of limited available previously published work (25,32–41) and local expert opinion when unavailable in the literature. The baseline utility was assumed to be 1 for the purposes of this study as the reference patient is assumed to be asymptomatic with no significant comorbidities, and there is no expected difference in baseline utility between the two strategies. Transient quality of life decrements that are not expected to last more than 3 months (such as FNA-related anxiety/discomfort or short-term surgical complications) were not included due to the challenges of incorporating these into a decision tree format and they were not felt to have a significant impact on the calculated end QALYs.

Data Analysis

The statistical analyses were performed in accordance with the Consolidated Health Economic Evaluation Reporting Standards (CHEERS) statement (42) and Drummond et al (43). The base case costs and QALYs (under the previously noted assumptions) were determined for each of the ATA2015 and TIRADS guidelines, with the difference in costs divided by the difference in QALYs to calculate the incremental cost-effectiveness ratio (ICER) and which was the optimal strategy. A willingness to pay (WTP) threshold was assumed at $50,000 per QALY as is commonly cited in the Canadian context (44), with evaluation of a $30,000 per QALY threshold considered as an alternative scenario in sensitivity analyses as this has been estimated as an implicit WTP threshold by Claxton et al (45).

Two additional sets of alternate scenarios were also evaluated: The first set allowed for a more generous surveillance schedule of ultrasounds being performed every year for 5 years for (1) all nodules meeting follow-up criteria under the ATA2015 strategy or (2) only the high suspicion nodules under ATA2015. Both scenarios were not uncommon at our institution *(M.Dahl, personal communication, Sept 2022)* prior to the release of the TIRADS guidelines given the lack of guidance regarding surveillance intensity and duration in the ATA2015 publication. The second set of alternate scenarios examine the effect of discounting at 0% and 3%, as recommended by CADTH guidelines (14). Distributional and heterogeneity concerns are not addressed with this model as no subgroup analyses were performed.

To explore the impact of uncertainty in the model, a series of univariate sensitivity analyses was performed and depicted as a Tornado diagram to examine how variation in parameters impacted the model and which were the most influential. Two-way sensitivity analyses were also performed for select parameters. The parameter ranges for sensitivity analysis were derived from previously published literature and observational data where available, with the maximum literature reported/observational ranges informing the upper and lower bounds. For parameters where there was insufficient data: (1) costs ranges were assumed at 50% and 150% of the base case estimate (similar to the methodology of Venkatesh et al (36) to reflect wide variability in practice and patient/disease course, including the potential for increased costs of complications and downstream surgery); (2) probabilities were varied by up to +/− 20% of the base case estimate.

Probability sensitivity analysis was performed using a Monte Carlo simulation of 10,000 trials. Probability distributions were assigned to each parameter as detailed in Table 2, specifically beta distributions for binomial probability and utilities, Dirichlet for multinomial probabilities, and gamma for costs (43). The 5-year probability of metastases from untreated thyroid cancer and survival of untreated thyroid cancer were assumed as fixed values as the former is essentially 0 and the latter is based on large scale population data (46). Where there was insufficient published data to

TABLE 2. Model Input Parameters

Parameters	Base Estimate	Ranges for Sensitivity Analysis	Probability Distribution	Source
Probabilities				
TIRADS				
Nodule requiring no follow-up	0.345	0.250–0.440	Dirichlet (34.5;23.7;41.8)	Pooled Institutional data and Hoang et al.(49)
Nodule requiring surveillance	0.237	0.205–0.269		
Nodule requiring FNA	0.418	0.291–0.545		
Nodule assigned to surveillance classified as:				
TIRADS 3	0.360	0.288–0.432	Dirichlet (36.0;54.4;9.6)	Pooled Institutional data and Hoang et al(49) for base case. Sensitivity analysis ranges +/- 20%.
TIRADS 4	0.544	0.435–0.653		
TIRADS 5	0.096	0.077–0.115		
Nodule assigned to surveillance upgraded to FNA threshold during surveillance period	0.139	0.051–0.154	Beta (167,1025)	Pooled from institutional data, Wong et al, Durante et al(25,53)
Nodule assigned to FNA returning malignant cytology	0.160	0.080–0.259	Beta (376,2147)	Pooled data from(20,21,54,55)
Missed malignancy in a nodule that was not followed-up	0.034	0.008–0.034	Beta (39,1031)	Middleton et al(20), Koseoglu et al(56)
ATA2015				
Nodule requiring no follow-up	0.027	0.019–0.035	Dirichlet (2.7;32.2;65.1)	Pooled Institutional data and Hoang et al(49).
Nodule requiring surveillance	0.322	0.250–0.394		
Nodule requiring FNA	0.651	0.587–0.715		
Nodule assigned to surveillance classified as:				
High suspicion	0.078	0.062–0.094	Dirichlet (7.8;81.0;11.2)	Pooled Institutional data and Hoang et al(49) for base case.
Low/Intermediate suspicion	0.810	0.658–0.972		
Very low suspicion	0.112	0.09–0.134		
Nodule assigned to surveillance upgraded to FNA threshold during surveillance period	0.135	0.02–0.154	Beta (163,1038)	Pooled from institutional data, Wong et al, Durante et al(25,53)
Nodule assigned to FNA returning malignant cytology	0.104	0.065–0.147	Beta (450, 3941)	Pooled data from(20,21,54,55)
Missed malignancy in a nodule that was not followed-up	0.038	0.030–0.046	Beta (19, 452)	Middleton et al(20). Sensitivity analysis range +/- 20%.
Other				
Long-term hypothyroidism after hemithyroidectomy	0.220	0.190–0.270	Beta (94,334)	Verloop et al(57)
Long-term unilateral RLN injury after hemithyroidectomy	0.010	0.009–0.011	Beta (396, 39203)	Leiker et al(38) Najafzadeh et al(41)
5-year survival untreated thyroid cancer	0.981	Fixed	Fixed	Davies and Welch(46)
Risk of distant metastatic disease in untreated thyroid cancer	0.0004	Fixed	Fixed	Saravana-Bawan et al.(58)

TABLE 2 (Continued)

Parameters	Base Estimate	Ranges for Sensitivity Analysis	Probability Distribution	Source
Costs				
Diagnostic thyroid ultrasound	$69	$35–$173	Gamma (16, 4.3)	See Appendix
Thyroid FNA	$241	$121–$603	Gamma (16,15)	
Hemithyroidectomy	$6375	$3188–$15938	Gamma (16, 399)	
Thyroid hormone supplementation (5 years)	$195	$98–$488	Gamma (16,12.2)	
Management of RLN injury	$11398	$5699–$28495	Gamma (16, 712)	
Utilities				
Benign cytology or subclinical disease	1.000	0.990–1.000	Beta (200,1)	Expert opinion
Stable nodule on surveillance	0.986	0.971–1.000	Beta (71.5, 1.0)	Wong et al(26), Kuo et al(33)
Thyroid cancer metastases	0.250	0.250–0.825	Beta (5.9, 5.1)	White et al(34), Houten et al(32)
Thyroid cancer mortality	0.000	Fixed	Fixed	Definition
Uncomplicated hemithyroidectomy	0.990	0.890–1.000	Beta (391, 3.9)	Esnaola et al(35), Venkatesh et al(36), Lang et al(37).
Long-term hypothyroidism	0.910	0.830–0.990	Beta (23.7, 2.4)	Base case averaged from Leiker et al(38), Heller et al(39), White et al(34)
Long-term RLN injury	0.627	0.560–0.690	Beta (149, 88)	Kebewew et al(40), Najafzadeh et al(41)

FNA, fine needle aspiration; RLN, recurrent laryngeal nerve. Dirichlet distributions specified by a list of alpha parameters; beta distributions specified by the parameters (alpha, beta); gamma distribution by the parameters (shape, scale). Costs are per unit, and further detailed in the appendix.

determine distribution parameters: Gamma distributions (costs): the reference value was considered the mean, and 50% of the reference value was assumed to represent 2 standard deviations; Beta distributions: The reference value was considered the mean, while 10% of the reference value was considered 1 standard deviation.

RESULTS

The model input parameters are summarized in Table 2.

Base Case and Scenarios

In the base-case scenario, TIRADS dominated the ATA2015 strategy by a small margin. Both produced near-equivalent QALYs with ATA2015 at a slightly increased cost compared to TIRADS (4.823 QALYs at $659 versus 4.828 QALYs at $634 respectively). The results of the alternative scenarios yield similar results, with TIRADS remaining dominant when the discount rate is varied at 0% and 3%, as well as when the more intense surveillance schedules for ATA2015 nodules were considered (Table 3).

Deterministic Sensitivity Analysis

A tornado diagram in Figure 2 illustrates the results of multiple univariate sensitivity analyses, showing the top 10 most influential parameters.

The top two most influential parameters are the probabilities of a malignant result with FNA under the TIRADS or ATA2015 strategies; with TIRADS no longer cost-effective if greater than 20.6% of FNA biopsies under TIRADS or less than 7.4% under ATA2015 yield malignancy. If the WTP threshold is reduced to $30,000, TIRADS is no longer cost-effective if the previously mentioned values are at 19.7% and 8.0% respectively. As these two parameters are influential in opposing directions but are expected to trend in the same direction (i.e. if the population has a higher background incidence of thyroid cancer, the malignant biopsy rate of ATA2015 and TIRADS would both increase), a two-way sensitivity analysis was performed to assess how these two parameters co-varied, demonstrating that at a WTP threshold of $50,000, TIRADS remains cost-effective even at the previously determined threshold of a 20.6% TIRADS FNA malignancy rate of biopsy if the ATA2015 malignancy rate increases to greater than 10.3% or conversely, if the latter is at 7.4% but the former decreases below 16.0%.

The utilities of surveillance and benign nodule/subclinical disease were the third and sixth most influential variables respectively, with the former approaching but not reaching an incremental net monetary benefit (NMB) of 0 at its maximum value of 1. Given the small difference in QALYs gained in the base-case scenario and the majority of the QALY contributions from these two utility values, a two-way sensitivity analysis was also performed (Fig 4), which showed that if the utility of a benign nodule/subclinical disease is reduced to 0.99 instead of the base case estimate of 1.00, TIRADS is no longer

TABLE 3. Base Case and Scenario Analyses

Strategy	Cost (CAD$)	Incremental Cost	Effectiveness (QALY)	Incremental Effectiveness	ICER
Base Case Analysis					
TIRADS	634,84		4.828		
ATA2015	659.46	24.62	4.823	-0.005	Dominated
Discount rate 0%					
TIRADS	638.08		4.973		
ATA2015	661.06	22.98	4.968	-0.005	Dominated
Discount rate 3%					
TIRADS	631.79		4.692		
ATA2015	657.95	26.16	4.687	-0.005	Dominated
5-year annual ultrasound for high-suspicion nodules meeting surveillance criteria under ATA2015					
TIRADS	634.84		4.828		
ATA2015	665.65	31.01	4.823	-0.005	Dominated
5-year annual ultrasound for all nodules meeting surveillance criteria under ATA2015					
TIRADS	634.84		4.828		
ATA2015	738.87	104.03	4.823	-0.005	Dominated

ICER, incremental cost-effectiveness ratio.

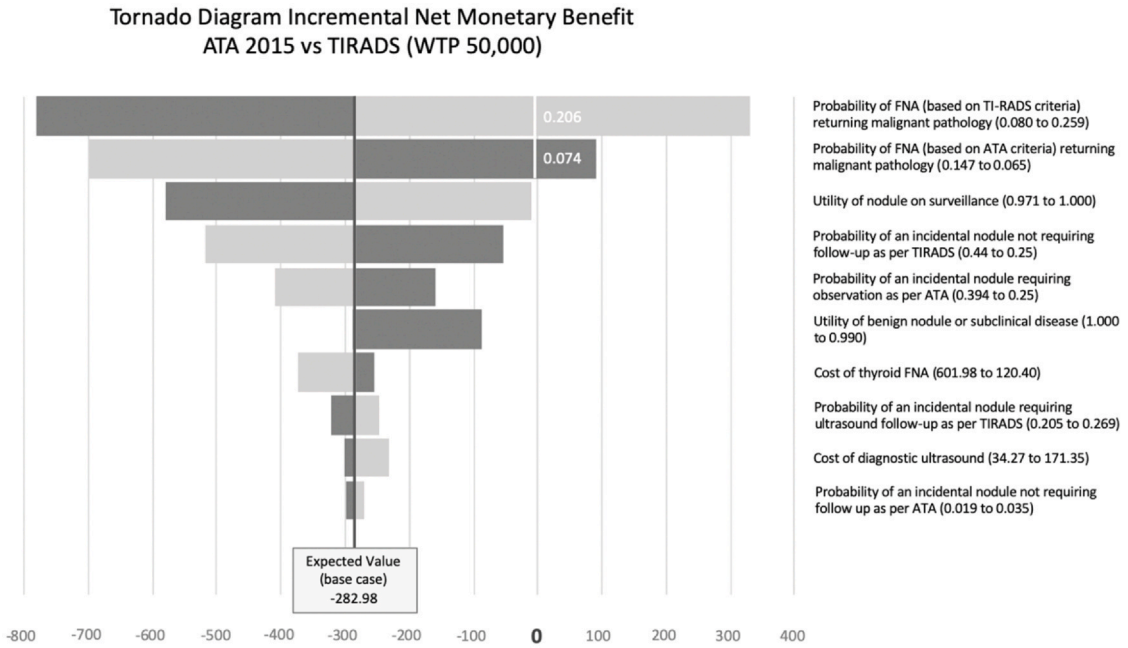


Figure 2. Incremental net monetary benefit tornado diagram (at a WTP threshold of $50,000) including the top 10 most influential parameters. The parameter descriptions on the right are followed by the sensitivity analysis ranges. A parameter with a dark gray bar on the left and light gray bar on the right indicates an increasing value as the incremental NMB increases, and conversely decreasing value when the colors are reversed. The white numbers on the top two bars indicate the values of those variables at which the net monetary benefit equates to 0 (i.e. TIRADS is no longer cost-effective). TIRADS, thyroid imaging reporting and data system; WTP, willingness-to-pay.

cost-effective if the utility of surveillance is higher than 0.99 at a WTP threshold of $50,000.

Probabilistic Sensitivity Analysis

PSA utilizing a Monte Carlo simulation showed that TIRADS was more cost-effective than ATA2015 in 79.7% of simulations at a WTP threshold of $50,000 per QALY (Fig 5). The majority of the scatterplot points are located in the north-west quadrant, indicating that ATA2015 is less effective with higher costs (i.e. TIRADS is dominant) in most of the simulation trials. The cost-effectiveness acceptability curve (Fig 6) also shows that TIRADS is favored in over 60% of simulations at all WTP thresholds from $0 to $100,000.

Stakeholder Engagement

This study greatly benefited from multidisciplinary input from endocrine surgery and endocrinology specialists (see

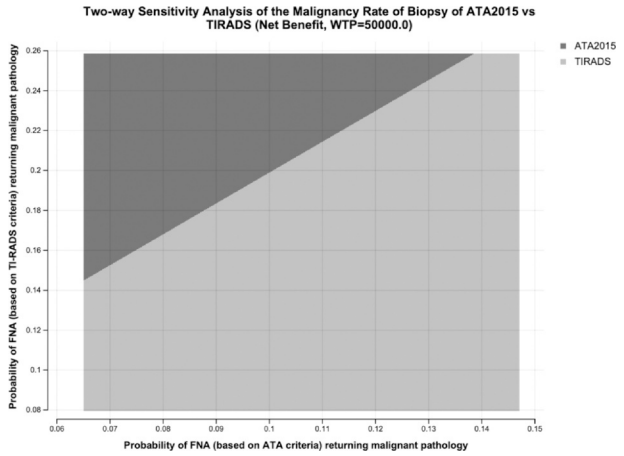


Figure 3. Results of two-way sensitivity analysis varying the probability of FNA yielding malignant pathology under TIRADS and ATA2015. The dark gray and light gray regions indicate where ATA2015 or TIRADS respectively would be more cost-effective (assuming a WTP threshold of $50,000) at a given combination of values for the probability of malignant FNA result under each strategy. The graph indicates that TIRADS is more cost-effective over most of the combinations of values within the sensitivity ranges. ATA, American Thyroid Association; FNA, fine needle aspiration; TIRADS, thyroid imaging reporting and data system.

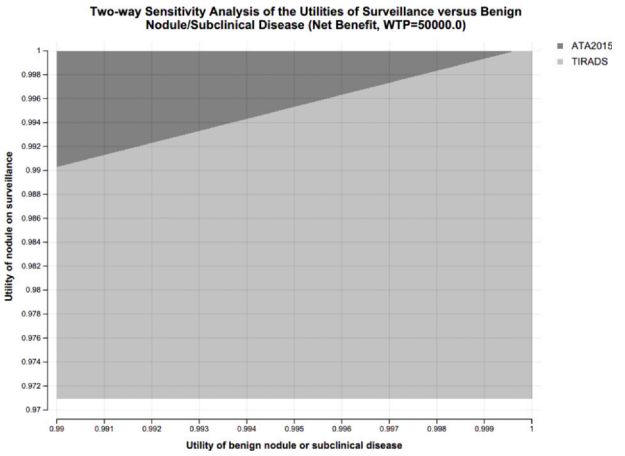


Figure 4. Results of two-way sensitivity analysis varying the utilities of a patient with a nodule under surveillance versus a known benign nodule or subclinical disease. The dark gray and light gray regions indicate where ATA2015 or TIRADS would be more cost-effective (assuming a WTP threshold of $50,000), respectively. The graph indicates that TIRADS is more cost-effective over most of the combinations of values within the sensitivity ranges. ATA, American Thyroid Association; TIRADS, thyroid imaging reporting and data system; WTP, Willingness-to-pay.

Acknowledgments), particularly in helping to streamline the surgery-related decision nodes in the model and providing historical experience with thyroid nodule surveillance practices prior to TIRADS being published.

DISCUSSION

Results Interpretation

The study analyses favor the TIRADS strategy being more cost-effective than ATA2015 by a small margin in the majority of cases. This remains the case when considering both ends of the spectrum of potential surveillance schedules under ATA2015, when the discount rate is varied from 0 to 3%, and at WTP thresholds of $30,000 and $50,000.

Although the model was found to be most sensitive to the rate of malignant biopsy under both strategies, they are influential in opposing directions but would be expected to trend in the same direction (as explained in the results section), thus lessening the individual impact of either variable as seen with the two-way sensitivity analyses.

The model was also sensitive to the estimated utilities of surveillance and subclinical disease or a known benign nodule. The less utility decrement experienced by a patient under surveillance or the more utility decrement experienced by the patient with a benign nodule/subclinical disease (perhaps due to anxiety), the less likely it is that TIRADS would be cost-effective. Given the small differences in incremental effectiveness between the two strategies, even small changes in either or both utility estimates could have a significant impact on the ICER as shown in the two-way sensitivity analyses. This relates to a major limitation of this study—namely the limited quality of life literature available to reference—specifically there was a paucity of studies which use standardized quality of life assessment tools (such as the EuroQoL EQ-5D) in evaluating thyroid nodule outcomes and post-operative complications specific to hemithyroidectomy (32). The only relevant study utilizing such a tool at the time of writing was Wong et al (26) who estimated a mean utility decrement of 0.029 at 18 months of surveillance using EQ-5D-5 L, however, no longer term measurements were performed. Kebebew et al used a standard gamble method to estimate a limited number of surgical utilities (40). Most other primary sources appear to rely on expert opinion (33,35,39). This apparent gap in the literature would benefit from future study with further patient engagement.

We initially hypothesized that there would be significantly fewer surgeries performed under the TIRADS strategy during the 5-year time horizon as more nodules undergo surveillance rather than biopsy. However, we found with rollback analysis of the base-case model that the probability of undergoing surgery for a malignant biopsy result was actually similar for both, 7.3% for ATA2015 versus 7.1% for TIRADS. This is likely because while fewer nodules meet FNA criteria under TIRADS compared to ATA2015, a higher percentage of these nodules are likely to yield a malignant result upon biopsy, consistent with other studies which have found that TIRADS has improved specificity for malignancy compared to ATA2015 (47). With the surgical costs being similar, the model suggests that the costs of increased surveillance in TIRADS are negated by the costs of increased benign biopsies in ATA2015 with similar effectiveness.

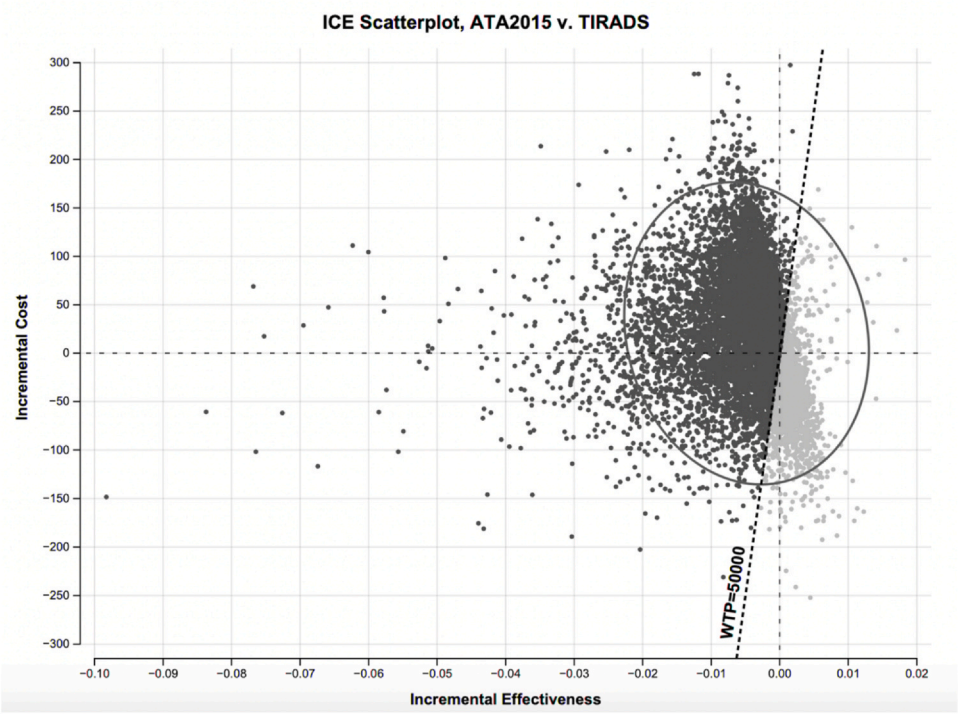


Figure 5. Results of Monte Carlo probabilistic sensitivity analysis of 10,000 trials in scatterplot format of ATA2015 vs TIRADS at a WTP threshold of $50,000 per QALY (black dashed line). The majority of the points are located in the northwest quadrant, indicating less effectiveness and higher costs associated with ATA2015 (i.e. TIRADS is dominant) in most simulations. The ellipse encompasses the 95% confidence interval. ATA, American Thyroid Association; QALY, quality adjusted life year; TIRADS, thyroid imaging reporting and data system; WTP, Willingness-to-pay.

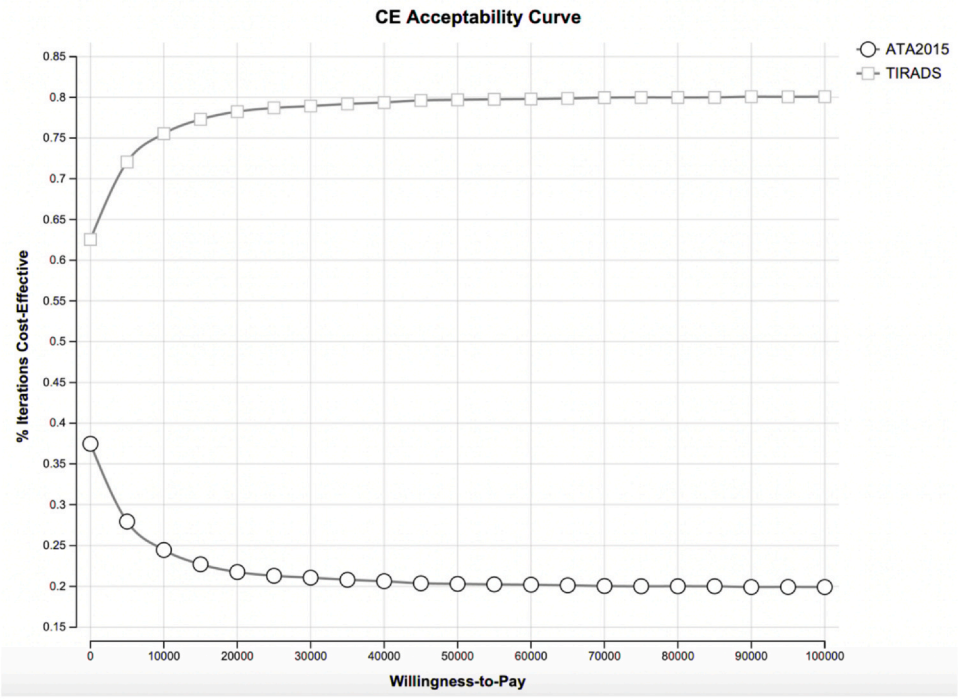


Figure 6. Cost-effectiveness acceptability curve of ATA2015 vs TIRADS. ATA, American Thyroid Association; TIRADS, thyroid imaging reporting and data system.

Assumptions and Related Limitations

In general terms, any model-based economic analysis not directly tied to a clinical trial will necessarily require assumptions based on the available evidence with inherent uncertainty, yet modeling remains the most common approach in cost-effectiveness studies due to practicality (48). Although a model cannot replicate what occurs in the real world exactly, it can give an idea of what may occur in common or plausible clinical scenarios. Full transparency regarding the model assumptions and a robust sensitivity analysis to test the model across a reasonable range of possibilities are necessary. In the case of our study, there were several factors underlying the rationalization for the model assumptions, the first being that a decision analysis model comparing these two strategies is to our knowledge, without published precedent and thus there are no established models to build upon.

Second, there is a scarcity of literature specifically evaluating incidentally discovered thyroid nodules that are risk-stratified using the ATA2015 and TIRADS guidelines, including the natural history of nodules which undergo surveillance under either strategy. One potential reason for this may be because the TIRADS guidelines were published relatively recently (in 2017) and may have taken some time to disseminate. Most studies comparing the two guidelines evaluate nodules that have already been biopsied, introducing significant sampling bias and likely overestimating the malignancy rate. The study by Hoang et al (included in the model inputs) prospectively included unselected thyroid nodules from multiple institutions and was the closest approximation of the intended study population (exclusion criteria included the history of thyroid cancer and nodules > 5 cm), however, it is unknown precisely what proportion of nodules in the study were incidentally discovered on imaging (49). It was for this reason that we supplemented the study with data from our own institution. These relative proportions would also be expected to vary with each jurisdiction depending on the referral base and population incidence of thyroid cancer.

Thirdly, there is significant variability in practice and many factors which impact the management of thyroid nodules in terms of surveillance, decision to perform repeat biopsies, and cancer treatment that it would be impractical to model all possibilities. Some of the more notable exclusions in this model for these reasons include nodules for which there is a nondiagnostic or indeterminate result, where a repeat biopsy is indicated due to concern for a false negative result, molecular testing, and patients who initially undergo total thyroidectomy or require completion thyroidectomy after initial hemithyroidectomy. The model presented in this study represents some of the more common pathways, and we hope that future work can examine these additional scenarios.

Regarding nondiagnostic and indeterminate FNA, it is interesting to note that indeterminate (Bethesda III/IV) nodules have a significant rate of diagnostic hemithyroidectomy (50) despite a malignancy rate of only 5–30% (9). Surgery is also recommended as an option for nodules with repeatedly non-diagnostic cytology (Bethesda I), with expected malignancy rate $< 5\%$ (9,51). The likelihood is that more absolute numbers of nondiagnostic and indeterminate nodules would be detected with ATA2015 as more patients undergo upfront biopsy with this strategy compared to TIRADS, thus one could speculate that inclusion of Bethesda I/III/IV nodules might result in increased surgical costs for ATA2015 compared to TIRADS with little clinical or quality of life benefit.

Other Limitations and Future Directions

The 5-year time horizon of the study was also a limitation. While this was predicated on the 5-year maximum surveillance interval for nodules under the TIRADS strategy and assumed under the ATA2015 strategy, it omits the additional years of surveillance that may be required for nodules which are upgraded during surveillance (but remain beneath FNA threshold) and would not take into account any interventions that might occur afterward (for example, if a nodule on surveillance was found to be malignant at the end of the 5 years and required surgery subsequently) or any disease progression that may occur after 5 years in malignant nodules which were not followed-up. However, the probability of a nodule under surveillance being found to be malignant was less than 1% under both strategies and the probability of death or metastases in missed malignancy was essentially 0%, thus even with these limitations it is unlikely that a longer time horizon would change the results substantially. This could be assessed in future models which could utilize a 10-year or even lifetime horizon when more data are available.

Another limitation was that short-term costs and effects were not considered, including post-operative hematoma or utility decrement associated with FNA. However, the risk of the former is very low (1%) and the latter is generally transient, with Wong et al showing no significant difference in utility decrement after 3 months between patients undergoing FNA and surveillance (26).

The model perspective from a publicly funded healthcare system in Canada may not be generalizable to other settings. Although a wide range of costs was considered in the sensitivity analyses, there may be settings where healthcare costs may be outside the range of those used in the study including privatized healthcare systems or developing economies. Most of the literature cited to estimate the model parameters were from American or Canadian institutions (where TIRADS is most in use), thus the model may not be as applicable to other geographic regions. Of note, the WTP threshold of $50,000 has historically been applied also in American

settings (currency conversion notwithstanding) though it has come under scrutiny for possibly being too low (52). The study is also limited to a single provincial single-payer perspective, which would be most useful to inform regional standardization of thyroid nodule management in BC, however, a wider Canadian or societal perspective (the latter which could include impacts on work/home productivity and patient borne costs) would be of interest as an avenue for future study. This study has also helped to identify other gaps in the literature such as quality of life data relating to thyroid nodule management and studies specifically evaluating incidentally discovered nodules despite this being a common clinical scenario.

The results of this study suggest that the utilization of TIRADS does not pose an economic disadvantage compared to ATA2015, and may be preferred for those who find the clarity of management recommendations and relative ease of use of TIRADS advantageous over ATA2015. It may also help to refute some of the concerns regarding the increased volume of surveillance and potential adverse outcomes from missed malignancies under TIRADS. There are currently efforts underway to incorporate TIRADS with other RSSs and develop unified international guidelines (17), and it is our hope that this study will help encourage a greater economic consideration.

FUNDING

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. The authors declare no conflict of interest.

DECLARATION OF COMPETING INTEREST

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

ACKNOWLEDGMENTS

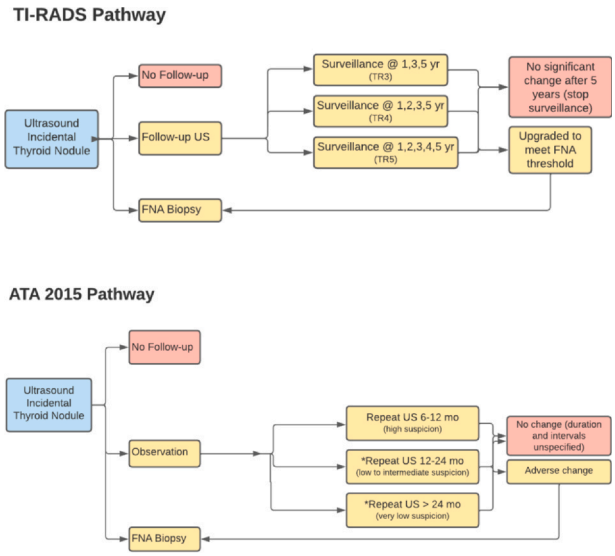
We gratefully acknowledge Drs T. Milner, S. Turkdogan, and M. Dahl for their time and expert opinion as well as Dr. E. Khalifa for assistance with data collection.

APPENDIX

A.1.1: TIRADS vs ATA2015 Clinical Algorithms

Graphical representation of the clinical pathways for TIRADS and ATA2015. ATA2015 discusses how to manage nodules after biopsy but this is outside the scope of the TIRADS white paper, thus downstream steps after FNA are omitted for clarity of comparison.

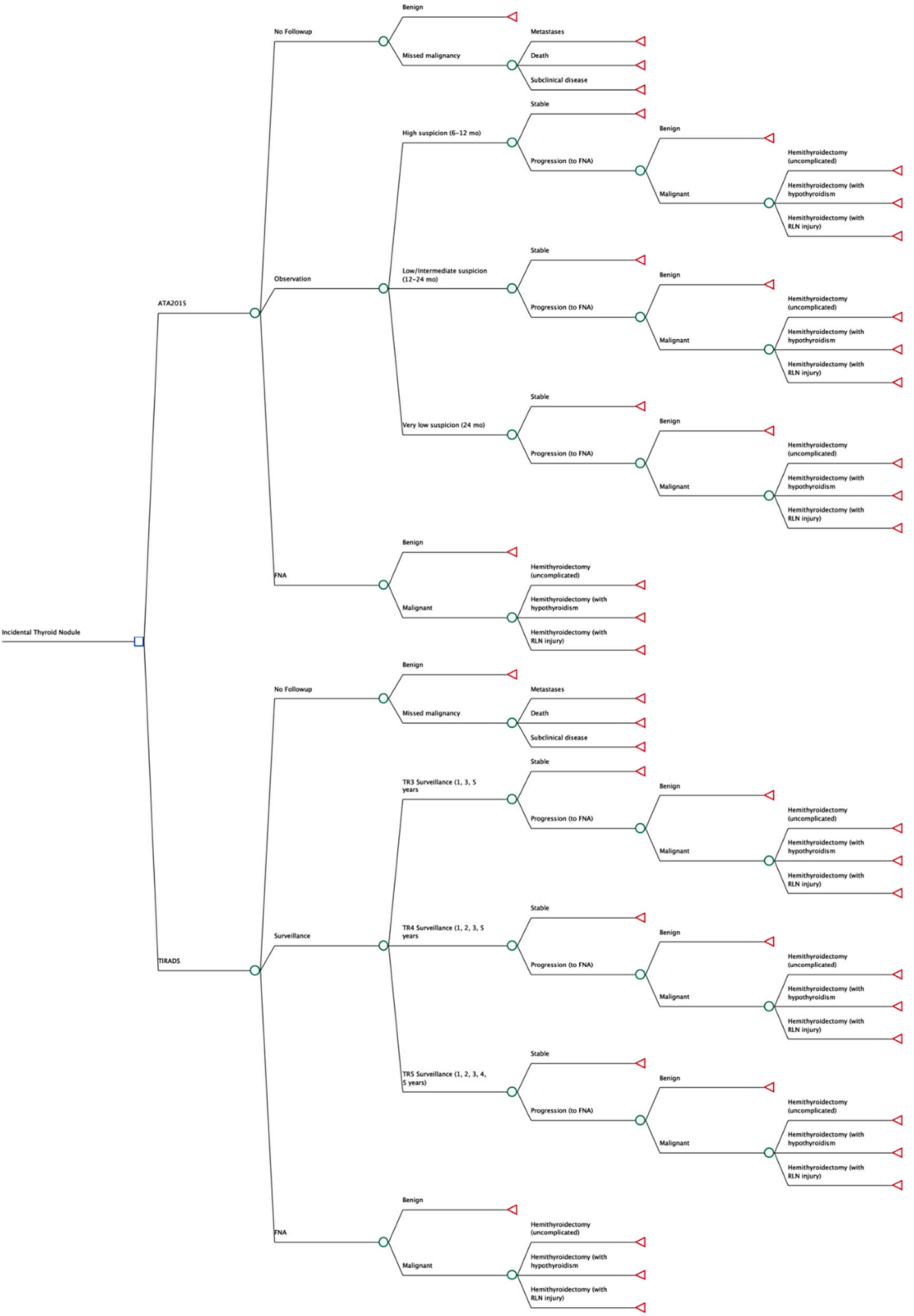
FNA, fine needle aspiration; US, ultrasound



.

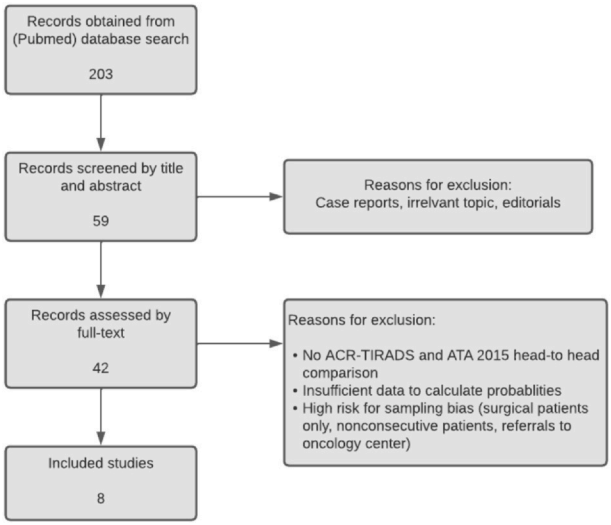
A.1.2: Detailed Decision Tree Model

FNA, fine needle aspiration; US, ultrasound; RLN, recurrent laryngeal nerve.


Incidental Thyroid Nodule
ATA2015
No Followup
Benign
Missed malignancy
Metastases
Death
Subclinical disease
Observation
High suspicion (6–12 mo)
Stable
Progression (to FNA)
Benign
Malignant
Hemithyroidectomy (uncomplicated)
Hemithyroidectomy (with hypothyroidism
Hemithyroidectomy (with RLN injury)
Low/Intermediate suspicion (12–24 mo)
Very low suspicion (24 mo)
FNA
TIRADS
Surveillance
TR3 Surveillance (1, 3, 5 years
TR4 Surveillance (1, 2, 3, 5 years
TR5 Surveillance (1, 2, 3, 4, 5 years)

A.2: PRISMA diagram of systematic literature search results

Used to populate model inputs related to TIRADS vs ATA2015 nodule risk stratification.



A.3. Detailed Accounting of Model Input Costs

Item	Unit Cost (in 2021 Canadian Dollars)	Assumptions and References
Thyroid ultrasound	$68.54	MSP 08642(30)
Thyroid FNA		
Diagnostic ultrasound fee	$68.54	MSP 08642(30)
Procedural fee	$72.79	MSP S00744(30)
Mini tray fee	$5.22	MSP 00044(30)
Cytology	$94.24	OLS 93085(31)
Uncomplicated hemithyroidectomy		
Physician Costs		
Surgeon pre-operative consult	$108.85	MSP 02513(30)
Surgeon pre-operative laryngoscopy	$44.30	MSP 00909(30)
Anesthesia pre-operative consult	$60.85	MSP 01151(30)
Surgeon cost for hemithyroidectomy	$587.84	MSP V70742(30)
Anesthesia cost for hemithyroidectomy	$146.56	MSP 01174(30) (assuming 60 min, level 4 procedure)
Surgeon post-operative appointments	$209.7	MSP 02507(30) (assume 6x follow-up appointments in 5 years, first 2x every 6 months, then annually)
Surgeon laryngoscopy at first post-operative appointment.	$44.30	MSP 00909(30)
Hospital Costs		
Thyroid surgery hospital costs	$4682.00	CIHI(28,29), CMG 424, British Columbia, 2019-2020
Diagnostic/Lab Costs		
Surgical Pathology (thyroid lobectomy specimen)	$103.20	L865(50) OHIP Fee Code*
Post-operative surveillance ultrasounds	$202.55	MSP 08642(30) (annually x 3 years)
Annual TSH	$48.04	OLS 92325(31) (annually for 5 years)
Annual Thyroglobulin	$135.38	OLS 92280(31) (annually for 5 years)
Long-term hypothyroidism		
Thyroid hormone supplementation	$194.82	Assume Synthroid 125 mcg once daily x 5 years(51)
Long-term recurrent laryngeal nerve injury		
Injection laryngoplasty	$637.82	MSP 02433(30), 2x every 6 months
Medialization laryngoplasty	$8839.95	MSP 02434 (surgeon fee) $637.88 MSP 01151 (anesthesia consult) $60.85

A.3 (Continued)

Item	Unit Cost (in 2021 Canadian Dollars)	Assumptions and References
Speech Therapy	$1920.00	MSP 01175 (anesthesia fee, level 4 for 90 min) $251.22(52) CIHI CMG 075(29) (hospital costs) $7890.50 Assume 12 sessions at $160/hr(52), fees based on WorkSafeBC (BC government occupational health agency) rates

Future costs discounted at 1.5% per annum.

CIHI, Canadian Institutes for Health Information; CMG, Case Mix Group; MSP, British Columbia Ministry of Health Medical Services Plan Billing Code (May 2021); OLS, British Columbia Ministry of Health Outpatient Laboratory Services Billing Code (Nov 2022); OHIP, Ontario Health Insurance Plan; TSH, thyroid stimulating hormone.

*Specific values for BC are unknown as it is incorporated into a salary structure, thus using OHIP (from another Canadian province) as surrogate data.

A4.1. Observational Data—Population Characteristics

Population Characteristics	
Total number of patients	200
Age [mean, (range)]	67.5 (31–96)
Gender (M:F)	59:141
Modality the incidental nodule was initially detected on:	
– X-ray	1
– Ultrasound	8
– Computed Tomography (CT)	180
– Magnetic Resonance Imaging (MRI)	8
– Unspecified	3
Patients lost to follow-up* or deceased before the end of year 5	39

*Of note, the 5-year follow-up period was inclusive of the COVID pandemic which may have impacted follow-up compliance.

A4.2. Observational Data—Summary of Results of Retrospective Re-scoring of Incidental Nodules

TIRADS		ATA2015	
No Follow-up	50 (25.0%)	No Follow-up	7 (3.5%)
Observation	41 (20.5%)	Observation	50 (25.0%)
TIRADS 3	15	Very low suspicion	10
TIRADS 4	22	Low/Intermediate suspicion	37
TIRADS 5	4	High suspicion	3
Upgraded to FNA threshold during surveillance	5	Upgraded to FNA threshold during surveillance	1
FNA	109 (54.5%)	FNA	143 (71.5%)
Total nodules = 200			

FNA, fine needle aspiration.

REFERENCES

1. Dean DS, Gharib H. Epidemiology of thyroid nodules. Best Pract Res Clin Endocrinol Metab 2008; 22(6):901–911. https://doi.org/10.1016/j.beem.2008.09.019
2. Papini E, Guglielmi R, Bianchini A, et al. Risk of malignancy in non-palpable thyroid nodules: predictive value of ultrasound and color-doppler features. J Clin Endocrinol Metab 2002; 87(5):1941–1946. https://doi.org/10.1210/jcem.87.5.8504
3. Yoon DY, Chang SK, Choi CS, et al. The prevalence and significance of incidental thyroid nodules identified on computed tomography. J Comput Assist Tomogr 2008; 32(5):810–815. https://doi.org/10.1097/RCT.0b013e318157fd38
4. Nam-Goong IS, Kim HY, Gong G, et al. Ultrasonography-guided fine-needle aspiration of thyroid incidentaloma: correlation with pathological findings. Clin Endocrinol (Oxf) 2004; 60(1):21–28. https://doi.org/10.1046/j.1365-2265.2003.01912.x
5. Smith-Bindman R, Lebda P, Feldstein VA, et al. Risk of thyroid cancer based on thyroid ultrasound imaging characteristics: results of a population-based study. JAMA Intern Med 2013; 173(19):1788–1796. https://doi.org/10.1001/jamainternmed.2013.9245
6. Hoang JK, Nguyen XV. Understanding the risks and harms of management of incidental thyroid nodules: a review. JAMA Otolaryngol Neck Surg 2017; 143(7):718. https://doi.org/10.1001/jamaoto.2017.0003
7. National Program of Cancer Registries and Surveillance, Epidemiology, and End Results Program. Cancer of the Thyroid - Cancer Stat Facts. SEER. Accessed December 5, 2022. ⟨https://seer.cancer.gov/statfacts/html/thyro.html⟩.
8. Vaccarella S, Franceschi S, Bray F, et al. Worldwide thyroid-cancer epidemic? The increasing impact of overdiagnosis. N Engl J Med 2016; 375(7):614–617. https://doi.org/10.1056/NEJMp1604412
9. 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):1–133. https://doi.org/10.1089/thy.2015.0020
10. 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(5):587–595. https://doi.org/10.1016/j.jacr.2017.01.046

11. Hoang JK, Asadollahi S, Durante C, et al. An international survey on utilization of five thyroid nodule risk stratification systems: a needs assessment with future implications. Thyroid 2022; 32(6):675–681. https://doi.org/10.1089/thy.2021.0558
12. Fradin JM. ACR TI-RADS: an advance in the management of thyroid nodules or Pandora's box of surveillance? J Clin Ultrasound 2020; 48(1):3–6. https://doi.org/10.1002/jcu.22772
13. Tessler FN, Middleton WD, Grant EG, et al. Re: ACR TI-RADS: an advance in the management of thyroid nodules or Pandora's box of surveillance? J Clin Ultrasound 2020; 48(1):7–8. https://doi.org/10.1002/jcu.22791
14. Guidelines for the Economic Evaluation of Health Technologies: Canada. 4th ed., CADTH,; 2017.
15. Pinchot SN, Al-Wagih H, Schaefer S, et al. Accuracy of fine-needle aspiration biopsy for predicting neoplasm or carcinoma in thyroid nodules 4 cm or larger. Arch Surg Chic Ill 1960 2009; 144(7):649–655. https://doi.org/10.1001/archsurg.2009.116
16. Pang Z, Margolis M, Menezes RJ, et al. Diagnostic performance of 2015 American Thyroid Association guidelines and inter-observer variability in assigning risk category. Eur J Radiol Open 2019; 6:122–127. https://doi.org/10.1016/j.ejro.2019.03.002
17. Hoang JK, Middleton WD, Tessler FN. Update on ACR TI-RADS: successes, challenges, and future directions, from the AJR special series on radiology reporting and data systems. Am J Roentgenol 2021; 216(3):570–578. https://doi.org/10.2214/AJR.20.24608
18. Cordes M, Götz TI, Horstrup K, et al. Growth rates of malignant and benign thyroid nodules in an ultrasound follow-up study: a retrospective cohort study. BMC Cancer 2019; 19(1):1139. https://doi.org/10.1186/s12885-019-6348-z
19. Cibas ES, Ali SZ. The 2017 Bethesda system for reporting thyroid cytopathology. Thyroid 2017; 27(11):1341–1346. https://doi.org/10.1089/thy.2017.0500
20. Middleton WD, Teefey SA, Reading CC, et al. Comparison of performance characteristics of American College of Radiology TI-RADS, Korean Society of Thyroid Radiology TIRADS, and American Thyroid Association Guidelines. Am J Roentgenol 2018; 210(5):1148–1154. https://doi.org/10.2214/AJR.17.18822
21. Koc AM, Adıbelli ZH, Erkul Z, et al. Comparison of diagnostic accuracy of ACR-TIRADS, American Thyroid Association (ATA), and EU-TIRADS guidelines in detecting thyroid malignancy. Eur J Radiol 2020; 133:109390https://doi.org/10.1016/j.ejrad.2020.109390
22. Ho AS, Sarti EE, Jain KS, et al. Malignancy rate in thyroid nodules classified as Bethesda Category III (AUS/FLUS). Thyroid 2014; 24(5):832–839. https://doi.org/10.1089/thy.2013.0317
23. Reuters KB, Mamone MCOC, Ikejiri ES, et al. Bethesda classification and cytohistological correlation of thyroid nodules in a brazilian thyroid disease center. Eur Thyroid J 2018; 7(3):133–138. https://doi.org/10.1159/000488104
24. Ullmann TM, Gray KD, Stefanova D, et al. The 2015 American Thyroid Association guidelines are associated with an increasing rate of hemithyroidectomy for thyroid cancer. Surgery 2019; 166(3):349–355. https://doi.org/10.1016/j.surg.2019.03.002
25. Wong CKH, Liu X, Lang BHH. Cost-effectiveness of fine-needle aspiration cytology (FNAC) and watchful observation for incidental thyroid nodules. J Endocrinol Invest 2020; 43(11):1645–1654. https://doi.org/10.1007/s40618-020-01254-0
26. Wong CKH, Lang BHH. A randomized trial comparing health-related quality-of-life and utility measures between routine fine-needle aspiration cytology (FNAC) and surveillance alone in patients with thyroid incidentaloma measuring 1–2 cm. Endocrine 2020; 67(2):397–405. https://doi.org/10.1007/s12020-019-02129-y
27. Bank of Canada. Bank of Canada Inflation Calculator. Accessed December 11, 2022. ⟨https://www.bankofcanada.ca/rates/related/inflation-calculator/⟩.
28. Philip Jacobs, Anthony Budden, Karen M.Lee. Guidance Document for the Costing of Health Care Resources in the Canadian Setting, 2nd ed. Published online March 2016. ⟨https://www.cadth.ca/sites/default/files/pdf/CP0009_CADTHCostingGuidance.pdf⟩.
29. Canadian Institute for Health Information (CIHI). CIHI Patient Cost Estimator. Accessed December 13, 2022. ⟨https://www.cihi.ca/en/patient-cost-estimator⟩.
30. British Columbia Ministry of Health. Medical Services Commission Payment Schedule. Published online May 1, 2021. Accessed December 13, 2022. ⟨https://www2.gov.bc.ca/assets/gov/health/practitioner-pro/medical-services-plan/msc-payment-schedule-may-2021.pdf⟩.
31. British Columbia Ministry of Health. Schedule of Fees For the Laboratory Services Outpatient Payment Schedule. Published online November 1, 2022. Accessed December 13, 2022. ⟨http://www.bccss.org/bcaplm-site/Documents/Programs/laboratory_services_schedule_of_fees.pdf⟩.
32. Houten R, Fleeman N, Kotas E, et al. A systematic review of health state utility values for thyroid cancer. Qual Life Res Int J Qual Life Asp Treat Care Rehabil 2021; 30(3):675–702. https://doi.org/10.1007/s11136-020-02676-2
33. Kuo EJ, Wu JX, Zanocco KA. Cost effectiveness of immediate biopsy versus surveillance of intermediate-suspicion thyroid nodules. Surgery 2018; 164(6):1330–1335. https://doi.org/10.1016/j.surg.2018.06.003
34. White C, Weinstein MC, Fingeret AL, et al. Is less more? A microsimulation model comparing cost-effectiveness of the revised American Thyroid Association's 2015 to 2009 guidelines for the management of patients with thyroid nodules and differentiated thyroid cancer. Ann Surg 2020; 271(4):765–773. https://doi.org/10.1097/SLA.0000000000003074
35. Esnaola NF, Cantor SB, Sherman SI, et al. Optimal treatment strategy in patients with papillary thyroid cancer: a decision analysis. Surgery 2001; 130(6):921–930. https://doi.org/10.1067/msy.2001.118370
36. Venkatesh S, Pasternak JD, Beninato T, et al. Cost-effectiveness of active surveillance versus hemithyroidectomy for micropapillary thyroid cancer. Surgery 2017; 161(1):116–126. https://doi.org/10.1016/j.surg.2016.06.076
37. Lang BHH, Wong CKH. Lobectomy is a more cost-effective option than total thyroidectomy for 1 to 4 cm papillary thyroid carcinoma that do not possess clinically recognizable high-risk features. Ann Surg Oncol 2016; 23(11):3641–3652. https://doi.org/10.1245/s10434-016-5280-6
38. Leiker AJ, Yen TW, Cheung K, et al. Cost analysis of thyroid lobectomy and intraoperative frozen section versus total thyroidectomy in patients with a cytologic diagnosis of "suspicious for papillary thyroid cancer. Surgery 2013; 154(6):1307–1314. https://doi.org/10.1016/j.surg.2013.06.031
39. Heller M, Zanocco K, Zydowicz S, et al. Cost-effectiveness analysis of repeat fine-needle aspiration for thyroid biopsies read as atypia of undetermined significance. Surgery 2012; 152(3):423–430. https://doi.org/10.1016/j.surg.2012.05.038
40. Kebebew E, Duh QY, Clark OH. Total thyroidectomy or thyroid lobectomy in patients with low-risk differentiated thyroid cancer: surgical decision analysis of a controversy using a mathematical model. World J Surg 2000; 24(11):1295–1302. https://doi.org/10.1007/s002680010215
41. Najafzadeh M, Marra CA, Lynd LD, et al. Cost-effectiveness of using a molecular diagnostic test to improve preoperative diagnosis of thyroid cancer. Value Health 2012; 15(8):1005–1013. https://doi.org/10.1016/j.jval.2012.06.017
42. Husereau D, Drummond M, Augustovski F, et al. Consolidated health economic evaluation reporting standards 2022 (CHEERS 2022) statement: updated reporting guidance for health economic evaluations. Value Health 2022; 25(1):3–9. https://doi.org/10.1016/j.jval.2021.11.1351
43. Michael FDrummond, Mark JSculpher, Karl Claxton, et al. Methods for the Economic Evaluation of Health Care Programmes. 4th ed., Oxford University Press,; 2015.
44. Griffiths E, Vadlamudi N. Cadth's $50,000 cost-effectiveness threshold: fact or fiction? Value Health 2016; 19(7):A488–A489. https://doi.org/10.1016/j.jval.2016.09.821
45. Karl Claxton. Estimating cost effectiveness thresholds for canadian provinces. Presented at: Network of Alberta Health Economics & Technology Assessment Rounds; September 2018. Accessed December 28, 2021. ⟨https://noahe.ca/events/past/noahe-rounds/health-rounds-ii/hetar-ii-s1⟩.
46. Davies L, Welch HG. Thyroid cancer survival in the United States: observational data from 1973 to 2005. Arch Otolaryngol Neck Surg 2010; 136(5):440. https://doi.org/10.1001/archoto.2010.55
47. Li W, Wang Y, Wen J, et al. Diagnostic performance of american college of radiology TI-RADS: a systematic review and meta-analysis. Am J Roentgenol 2021; 216(1):38–47. https://doi.org/10.2214/AJR.19.22691
48. Kadom N, Itri JN, Trofimova A, et al. Cost-effectiveness analysis: an overview of key concepts, recommendations, controversies, and pitfalls. Acad Radiol 2019; 26(4):534–541. https://doi.org/10.1016/j.acra.2018.10.014

49. Hoang JK, Middleton WD, Langer JE, et al. Comparison of thyroid risk categorization systems and fine-needle aspiration recommendations in a multi-institutional thyroid ultrasound registry. J Am Coll Radiol 2021; 18(12):1605–1613. https://doi.org/10.1016/j.jacr.2021.07.019
50. Kuta V, Forner D, Azzi J, et al. Treatment choices in managing bethesda III and IV thyroid nodules: a Canadian multi-institutional study. OTO Open 2021; 5(2):1–7. https://doi.org/10.1177/2473974×211015937
51. Espinosa De Ycaza AE, Lowe KM, Dean DS, et al. Risk of malignancy in thyroid nodules with non-diagnostic fine-needle aspiration: a retrospective cohort study. Thyroid 2016; 26(11):1598–1604. https://doi.org/10.1089/thy.2016.0096
52. Neumann PJ, Cohen JT, Weinstein MC. Updating cost-effectiveness — the curious resilience of the $50,000-per-QALY threshold. N Engl J Med 2014; 371(9):796–797. https://doi.org/10.1056/NEJMp1405158
53. Durante C, Costante G, Lucisano G, et al. The natural history of benign thyroid nodules. JAMA 2015; 313(9):926. https://doi.org/10.1001/jama.2015.0956
54. Pandya A, Caoili EM, Jawad-Makki F, et al. Retrospective cohort study of 1947 thyroid nodules: a comparison of the 2017 American College of Radiology TI-RADS and the 2015 american thyroid association classifications. Am J Roentgenol 2020; 214(4):900–906. https://doi.org/10.2214/AJR.19.21904
55. Merhav G, Zolotov S, Mahagneh A, et al. Validation of TIRADS ACR risk assessment of thyroid nodules in comparison to the ATA guidelines. J Clin Imaging Sci 2021; 11:37. https://doi.org/10.25259/JCIS_99_2021
56. Koseoglu Atilla FD, Ozgen Saydam B, Erarslan NA, et al. Does the ACR TI-RADS scoring allow us to safely avoid unnecessary thyroid biopsy? Single center analysis in a large cohort. Endocrine 2018; 61(3):398–402. https://doi.org/10.1007/s12020-018-1620-6
57. Verloop H, Louwerens M, Schoones JW, et al. Risk of hypothyroidism following hemithyroidectomy: systematic review and meta-analysis of prognostic studies. J Clin Endocrinol Metab 2012; 97(7):2243–2255. https://doi.org/10.1210/jc.2012-1063
58. Saravana-Bawan B, Bajwa A, Paterson J, McMullen T. Active surveillance of low-risk papillary thyroid cancer: a meta-analysis. Surgery 2020; 167(1):46–55. https://doi.org/10.1016/j.surg.2019.03.040