Ontario Health

ONTARIO HEALTH TECHNOLOGY ASSESSMENT SERIES

Molecular Testing for Thyroid Nodules of Indeterminate Cytology: A Health Technology Assessment

Key Messages

What Is This Health Technology Assessment About?

The thyroid is a gland in the lower neck that releases hormones related to growth and metabolism (the process of converting food into energy). Cancer in the thyroid gland can spread to other parts of the body, but not all thyroid nodules (growths) are cancerous. As well, some types of thyroid cancer are not aggressive and can be left alone.

Since the 1970s, the rate of thyroid cancer diagnoses has increased. Improved diagnostic accuracy with molecular testing could lead to better classification of thyroid nodules and result in fewer unnecessary treatments.

This health technology assessment looked at how accurate, useful, and cost-effective molecular testing is for people with thyroid nodules of indeterminate cytology (that is, the cells are not clearly cancerous but not clearly benign). It also looked at the budget impact of publicly funding molecular testing and at the experiences, preferences, and values of people with thyroid nodules of indeterminate cytology.

What Did This Health Technology Assessment Find?

Compared to usual care (no molecular testing), molecular testing for thyroid nodules of indeterminate cytology may be accurate as a "rule-out" test. Molecular testing might also lead to lower rates of surgery to remove nodules, but the evidence is uncertain.

Molecular testing is unlikely to be cost-effective at its current list price. Publicly funding molecular testing in Ontario over the next 5 years would cost an additional $6.24 million.

People with thyroid nodules valued the information they could get from molecular testing to help them make treatment decisions. They expressed concern about the time it takes to receive the results of molecular testing, especially if the findings are not conclusive or useful for decision-making.

Published April 2022 Volume 22, Number 2

Acknowledgments

This report was developed by a multidisciplinary team from Ontario Health. The clinical epidemiologist was Stacey Vandersluis, the medical librarian was Corinne Holubowich, the primary health economist was Chunmei Li, the secondary health economist was Xuanqian Xie, and the patient engagement analyst was David Wells.

The medical editor was Jeanne McKane. Others involved in the development and production of this report were Jamie Turnbull, Claude Soulodre, Susan Harrison, Elisabeth Smitko, Sarah McDowell, Olga Gajic-Veljanoski, Vivian Ng, Andrée Mitchell, Charles de Mestral, and Nancy Sikich.

We would like to thank the following people and organizations for lending their expertise to the development of this report:

- Lisa Caulley, University of Ottawa, Department of Otolaryngology–Head and Neck Surgery
- Antoine Eskander, Sunnybrook Health Sciences Centre and Michael Garron Hospital, Department of Otolaryngology–Head and Neck Surgery, Surgical Oncology
- Michael J. Odell, Department of Otolaryngology, Faculty of Medicine, University of Ottawa
- Katie O'Reilly, Department of Pathology Pathologist at Peterborough Regional Health Centre
- Aaron Pollett, Division of Diagnostic Medical Genetics, Mount Sinai Hospital
- Group TMTC (Thyroid Molecular Testing and Coordination)

We also thank our lived experience participants who generously gave their time to share their stories with us for this report.

The statements, conclusions, and views expressed in this report do not necessarily represent the views of those we consulted.

Parts of this health technology assessment are based on data and information provided by Ontario Health, and include data received from ICES, the Canadian Institute for Health Information (CIHI), and the Ministry of Health. The opinions, reviews, views, and conclusions reported in this publication are those of the authors and do not necessarily reflect those of Ontario Health, ICES, CIHI, and/or the Ministry of Health. No endorsement by Ontario Health, ICES, CIHI, and/or the Ministry of Health is intended or should be inferred.

Citation
Ontario Health. Molecular testing for thyroid nodules of indeterminate cytology: a health technology assessment. Ont Health Technol Assess Ser [Internet]. 2022 Apr;22(2):1–111. Available from: https://www.hqontario.ca/evidence-to-improve-care/health-technology-assessment/reviews-and-recommendations/molecular-testing-for-thyroid-nodules-of-indeterminate-cytology

Abstract

Background

The thyroid is a gland in the lower neck that is responsible for secreting hormones related to growth and metabolism. A cancer growth in the thyroid can spread to other parts of the body, but most thyroid nodules (growths) are benign, and some types of thyroid cancer are nonaggressive and can be managed with active surveillance only. We conducted a health technology assessment of molecular testing in people with thyroid nodules of indeterminate cytology, which included an evaluation of diagnostic accuracy, clinical utility, cost-effectiveness, the budget impact of publicly funding molecular testing, and patient preferences and values.

Methods

We performed a systematic literature search of the clinical evidence. We assessed the risk of bias of each included study using the Risk of Bias Among Systematic Review (ROBIS) tool for systematic reviews, the Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2) assessment for primary studies that evaluated diagnostic accuracy, and the Risk of Bias tool for Non-randomized Studies (RoBANS) for primary studies that evaluated clinical utility. We evaluated the quality of the body of evidence according to the Grading of Recommendations Assessment, Development, and Evaluation (GRADE) Working Group criteria. We performed a systematic economic literature review and conducted cost-effectiveness and cost–utility analyses with a 5-year time horizon from the Ontario Ministry of Health perspective. We also analyzed the budget impact of publicly funding molecular testing in people with thyroid nodules of indeterminate cytology in Ontario. To contextualize the potential value of molecular testing in people with thyroid nodules of indeterminate cytology, we spoke to people with thyroid nodules.

Results

In the clinical evidence review, we included one systematic review, which contained eight relevant primary studies. Using molecular testing to support the rule-out of cancer in thyroid nodules of indeterminate significance may reduce the number of unnecessary surgeries. For diagnostic accuracy, molecular testing for a diagnosis of malignancy in a nodule of indeterminate significance had a sensitivity of 91% to 94% and a specificity of 68% to 82% (GRADE: Low). As well, lower rates of surgical resections were reported in nodules of indeterminate cytology (GRADE: Very Low). Compared to diagnostic lobectomy, we found that molecular testing would increase the probability of predicting a correct diagnosis, reduce the probability of unnecessary surgery, and lead to a slight improvement in quality-adjusted life-years (QALYs), but it would increase costs. The resulting incremental cost-effectiveness ratio was $220,572 to $298,653 per QALY gained. At the commonly used willingness-to-pay values of $50,000 and $100,000 per QALY gained, molecular testing was unlikely to be cost-effective (probability of molecular testing being cost-effective was less than 50%). Publicly funding molecular testing in Ontario over the next 5 years would lead to an additional cost of $6.24 million. People with thyroid nodules of indeterminate cytology reported on the benefits and drawbacks of molecular testing, as well as barriers to accessing and choosing to undergo molecular testing.

Conclusions

For thyroid nodules of indeterminate cytology, molecular testing may have diagnostic accuracy as a rule-out test, and it may result in fewer nodule resections than usual care (no molecular testing). For people with thyroid nodules of indeterminate cytology, molecular testing at the current list price is unlikely to be cost-effective compared to diagnostic lobectomy. Publicly funding molecular testing in Ontario would cost about $6.24 million over the next 5 years. People with thyroid nodules of indeterminate cytology valued the information that could be provided by molecular testing, but they expressed concern about the time required to obtain results, especially if the findings were not conclusive or useful for treatment decision-making.

Table of Contents

List of Tables

List of Figures

Objective

This health technology assessment evaluates the clinical validity (diagnostic accuracy), clinical utility, and cost-effectiveness of molecular testing for people with thyroid nodules of indeterminate cytology. It also evaluates the budget impact of publicly funding molecular testing, and the experiences, preferences, and values of people with thyroid nodules, including those with indeterminate cytology.

Background

Health Condition

The thyroid is a gland in the lower neck that is responsible for secreting hormones related to growth and metabolism (the process of converting food into energy).[1] A cancer growth in the thyroid may spread to other parts of the body.[1] Previous radiation exposure—especially to the head and neck—is known to be a risk factor for thyroid cancer; as a result, some treatments that involved radiation (such as x-ray for acne) are no longer used.[2] Race/ethnicity is also a risk factor for thyroid cancer: people of Asian descent have higher incidence rates of thyroid cancer.[3]

There are four main types of thyroid cancer based on the type of thyroid cells they originate from: follicular, papillary, medullary, and anaplastic. Follicular and papillary cancers are the most common and most treatable types of thyroid cancer; medullary and anaplastic thyroid cancers are more rare and more difficult to treat. Observed diagnoses of papillary thyroid cancers have been increasing, and thus incidence rates have risen over time. However, papillary cancer is also typically nonaggressive, and some people can do well when it is managed with active surveillance only.[4]

Investigation of suspicious thyroid nodules (growths on the thyroid) typically begins with a clinical history and physical examination, which may include ultrasound imaging. Since the 1970s (around the time ultrasound imaging was introduced for thyroid nodules of concern), rates of thyroid cancer diagnoses and the identification of nodules of indeterminate cytology (i.e., according to the test results, it is unclear if the nodule is benign or malignant) have increased markedly.[4,5] Choosing Wisely Canada has recommended against the routine use of ultrasound in people with neck and throat pain or discomfort, or with abnormal thyroid function, if no palpable abnormality is found.[6,7] The use of ultrasound may inadvertently draw attention to nodules that are not causing problems,[6] supporting the theory that some thyroid nodules would likely not develop aggressive growth; this leads to potential concerns about over-diagnosis and over-treatment.[4] This theory is further supported by the fact that the mortality rate for thyroid cancers has remained stable over time (1975 to 2009).[8]

Improved diagnostic accuracy (i.e., classifying thyroid nodules as benign or malignant) may reduce the number of unnecessary invasive treatments such as surgery. There are multiple ways to improve the accurate classification of thyroid nodules, including the following: molecular testing; a multidisciplinary combination of approaches; uniform reporting of ultrasound characteristics; and quality assurance related to thyroid cytopathology (i.e., monitoring how often samples are identified as “atypical cells of undetermined significance”; Katie O’Reilly, email communication, July 21, 2020)

Clinical Need and Target Population

Thyroid nodules are quite common; they are identifiable in 19% to 68% of the general population.[9,10] However, most thyroid nodules are benign—only 7% to 15% are malignant.[9] In 2018, thyroid cancers made up 3.7% of all Ontario cancer diagnoses, at 3,341 cases.[11] Overall survival with thyroid cancers

is 98%, and death rates have remained stable over time, despite the fact that incidence rates have increased.[11,12]

Current Diagnostic Testing and Classification

First-line treatment for most thyroid cancers is surgery (thyroidectomy) to remove part or all of the affected thyroid, with or without the lymph nodes located beneath and around the gland (central neck dissection).[1] People who undergo thyroidectomy must be monitored closely long-term for impaired or lost thyroid function.[13] In 2017, the Cancer Care Ontario Thyroid Cancer Guideline group endorsed the 2015 American Thyroid Association management guidelines for thyroid cancer.[14,15] These guidelines specify that when clinically indicated, thyroid nodules should be investigated with ultrasound-guided fine-needle aspiration biopsy (removing a sample of cells using a thin needle and a syringe) and cytology (examining cells under a microscope), and then reported using The Bethesda System for Reporting Thyroid Cytopathology (TBSRTC).[9] Nodules that are potentially cancerous are classified using the TBSRTC, which uses six categories (from non-cancerous to malignant) to help guide diagnosis and treatment decisions. The TBSRTC was developed for pathologists[16]; the target risk of malignancy for each category is summarized in Table 1.

Table 1: The Bethesda Classification System and Risk of Malignancy, 2017

Diagnostic category and cytological diagnosis	Target risk of malignancy, %[a]	Usual clinical management
I: Nondiagnostic or unsatisfactory	5–10	Repeat fine-needle aspiration with ultrasound guidance
II: Benign	0–3	Clinical and sonographic follow-up
III: Atypical cell of undetermined significance/follicular lesion of undetermined significance	10–30	Repeat fine-needle aspiration, or conduct molecular testing or lobectomy
IV: Follicular neoplasm/suspicion for a follicular neoplasm	25–40	Molecular testing or lobectomy
V: Suspicious for malignancy	50–75	Near-total thyroidectomy or lobectomy
VI: Malignant	97–99	Near-total thyroidectomy or lobectomy

[a] For noninvasive follicular thyroid neoplasms with papillary-like nuclear features equivalent to carcinoma.
Source: Cibas et al.[17]

For this health technology assessment, we have followed the convention of considering nodules to be of indeterminate cytology if they fall into TBSRTC categories III or IV.[18] TBSRTC III (atypical cell of undetermined significance/follicular lesion of undetermined significance [AUS/FLUS]) was intended to have limited use and be used for less than 7% of all diagnoses.[17,19] However, between 2016 and 2019 in Ontario, 9.0% of cytology diagnoses were TBSRTC III and 2.8% were TBSRTC IV (based on an evaluation of hospital administrative databases; additional details in the primary economic analysis). This finding was comparable to rates reported by one Toronto hospital between 2010 and 2015.[20,21]

Cytology diagnosis around the AUS/FLUS category may involve high intra-observer variability, and a repeat fine-needle aspiration biopsy or a second opinion from an experienced pathologist might

improve accuracy.[9] In one study, 40% (17 of 43) of people with AUS/FLUS cytology findings who underwent repeat fine-needle aspiration were classified as AUS/FLUS again; the rest were re-classified into a different category.[22] As well, the predicted malignancy rates for TBSRTC III and IV (Table 1) may not be representative of reality for some institutions. For example, one Toronto hospital had malignancy rates of 36% for TBSRTC III and 55% for TBSRTC IV.[21] Other Ontario[23] and international estimates have reported malignancy rates of 6% to 48% (mean risk 16%) for nodules of indeterminate cytology, demonstrating substantial variability.[9,16]

The 2015 American Thyroid Association guidelines[9] recommend that nodules classified as TBSRTC III (AUS/FLUS) be actively monitored or surgically managed (partial or complete thyroidectomy). However, it has been proposed that surgical management is being overused; more than 70% of nodules of indeterminate cytology that were surgically removed were found to be benign with confirmatory histopathology.[24]

Health Technology Under Review

Molecular testing uses cells collected from fine-needle aspiration biopsy to assess the suspicious thyroid tissue sample at a genetic level. Often, after a first fine-needle aspiration biopsy for cytology and TBSRTC classification, a second one is required to obtain an adequate tissue sample for molecular testing. The introduction of molecular testing to improve diagnostic accuracy could support improved care and reduce unnecessary treatment. A proposed clinical pathway is outlined in Figure 1.

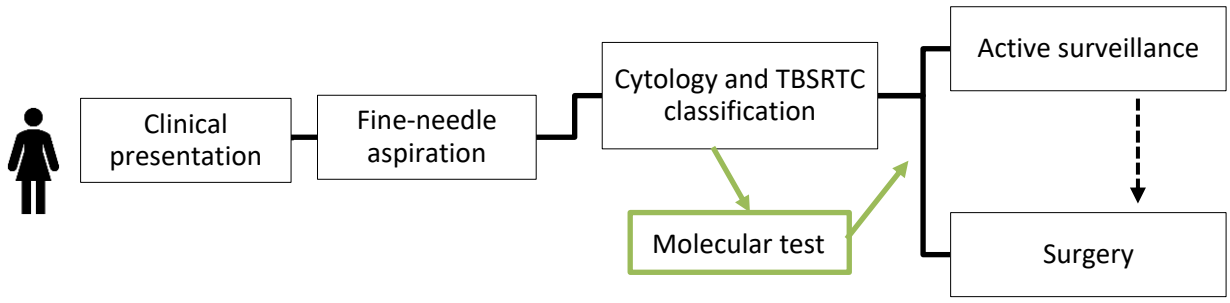

Figure 1: Proposed Modified Clinical Pathway With Molecular Testing

Abbreviation: TBSRTC, The Bethesda System for Reporting Thyroid Cytopathology.

The reference standard for diagnosing thyroid cancer is partial thyroidectomy and histopathological examination, with the potential for complete thyroidectomy as a second procedure.[25] As well, histopathological diagnoses use a binary system (benign or malignant), but this may minimize real variations in tumours, which present more like a spectrum.[25] For thyroid nodules of indeterminate cytology, the clinical benefit of a "rule-out" testing is that it may eliminate unnecessary surgeries (e.g., removal of what is ultimately a benign tumour). However, as the technology evolves and the evidence base for the molecular testing in a "rule-in" capacity develops, diagnostic testing may be considered in the future.[26]

Thyroid nodules have some well-established molecular features. The most common mutations in papillary thyroid cancer (approximately 80% of all thyroid cancers) are point mutations in the *BRAF*, *RAS*,

RET, and *NTRK* genes.[12,27] Follicular thyroid cancers are more typically characterized by mutations in the *RAS* and *PPARγ/PAX8* genes.[12,27] Mutations in the *MAPK* and *PI3K* genes have also been identified in thyroid cancers, because they are involved in the signalling pathway.[27] There is no consensus yet on how particular variants are used to identify characteristics of lesions and how they ought to be used to support treatment and management pathways.[28-31] The body of research into other genes and variants related to thyroid cancers is growing, adding to the complexity of this space and leading each available molecular test for thyroid cancer to have unique aspects.

Several molecular tests are commercially available, and these include anywhere from several dozen genes to hundreds. Although these tests share the same objective, the differences between them in terms of included molecular markers and techniques may lead to differences in diagnostic performance. Two tests have been used in Canadian research studies and are included in this health technology assessment: Afirma and ThyroSeq.[32,33] The Afirma gene expression classifier (GEC) test uses mRNA expression of 167 genes; the newer version is called the genomic sequence classifier (GSC).[34] The Afirma tests function as rule-out tests to confirm that a person is very unlikely to have cancer.[34] The ThyroSeq test (also known as the gene mutation panel [GMP]) is on its third version (ThyroSeq v3), introduced in 2017 as a 112-gene panel.[35] ThyroSeq uses genes and variants that enable it to be used as more of an all-around test for diagnosis (a rule-in and rule-out test).[25] Other similar molecular tests are available, but they did not meet the inclusion criteria for this health technology assessment. One of the tests that did not meet our criteria is the Afirma Xpression Atlas companion, which is intended to inform clinical decision-making (rule-in test).[26] Another is ThyGeNEXT with ThyraMIR, which together act as both rule-in and rule-out tests (the ThyGeNEXT test is a next-generation sequencing panel made up of 10 genes and 41 RNA variants; the ThyraMIR miRNA classifier is made up of 10 miRNA variants).[36]

Safety/Harms

There are no direct concerns about safety or harms related to molecular testing itself. The test is conducted using a tissue sample, known as fine-needle aspiration biopsy, and it can use cells already extracted for the cytology assessment, although a second fine-needle aspiration biopsy is usually required to obtain an adequate sample.

There may be harms related to false results because of an inaccurate test. A false positive result (i.e., a malignancy diagnosis) may lead to unnecessary treatment, such as the excision of an otherwise healthy thyroid. A false negative would miss potential cancer and lead to a missed opportunity for early diagnosis and treatment, which could increase a person's risk of mortality.

Regulatory Information

Health Canada licensing is not required for tests conducted out of country, and all known molecular tests for thyroid nodules of indeterminate cytology are conducted in the United States. Samples are collected in Canada but mailed out of country for testing.

Ontario, Canadian, and International Context

The level of need for molecular testing in Ontario is uncertain at present; dissemination pressure from patients and clinicians is low. A 2017 publication,[32] described as the first Canadian experience with Afirma, included patients from Montreal, Quebec, and St. John's, Newfoundland. In Ontario, physicians can offer people the option of paying out of pocket for the test (at a cost of $2,000 to $5,000), but very few people are willing undertake the test because of the cost (Lisa Caulley, email communication,

October 8, 2020). Fewer than five tests a year have been performed at high-volume centres (Lisa Caulley, email communication, October 8, 2020; Antoine Eskander, email communication, March 28, 2021; Mike Odell, email communication, April 6, 2021).

Cancer Care Ontario (now part of Ontario Health) has endorsed[14] the 2015 American Thyroid Association guidelines, which recommend that for nodules with AUS/FLUS cytology (TBSRTC III), molecular testing or repeat fine-needle aspiration may be used to supplement malignancy risk assessment before opting for surveillance or surgery (weak recommendation, moderate-quality evidence).[9] The guideline states that the benefits and limitations of molecular testing should be clearly communicated to patients when making a decision to test. It should also be noted that several guideline authors had a declared conflict of interest, because they received funding from Veracyte, the manufacturer of Afirma.

During the COVID-19 pandemic, people with well-differentiated thyroid cancers have been considered the lowest priority ("priority C") for surgery according to Ontario Health (Cancer Care Ontario), because a delay of 2 months would be unlikely to affect the outcome.[37] Such prioritization further underscores the potential role of molecular testing in helping to select those who would be most likely to benefit from thyroidectomy because of suspected cancer.[37] However, in practice molecular testing has not had much impact on the management of thyroid cancers in Ontario (Lisa Caulley, email communication, October 8, 2020).

In January 2021, the Quebec Institut national d'excellence en santé et en services sociaux (INESSS) made a recommendation[18] in favour of funding molecular tests, but with some restrictions suggested by a clinical expert advisory group. Specifically, INESSS recommended that molecular testing be used only for thyroid nodules that are classified as TBSRTC III after a second fine-needle aspiration, or for those classified as TBSRTC IV that are between 1 and 4 cm and who have intermediate (equivocal or discordant) clinical or sonographic risk factors. As well, INESSS recommended that testing be limited to circumstances when it would inform clinical management, and not for patients for whom surgery is indicated due to high risk factors, for those whom diagnostic lobectomy is not being considered due to low risk of malignancy, or for those who express a preference (for surgery or observation).[18] Furthermore, the INESSS recommendation is contingent on a validation study based on samples collected in Quebec to provide independent confirmation.[18]

In the United States, Afirma, ThyroSeq, and ThyGenX/ThyGeNEXT with ThyraMIR are covered by Medicare and many health insurance plans, including BlueCross/BlueShield.[38-40] Outside the United States, molecular testing has not been widely adopted. The authors of the 2017 European Thyroid Association guideline[6] convened a special panel of experts to discuss molecular testing and fine-needle aspiration cytology diagnostics; they concluded that there may be diagnostic value for mutation panels, for examining the evidence around specific genes, and for the Afirma GEC test. They also concluded that there may be benefit to considering genetic panels that include *BRAF, RET/PTC*, *PZX8/PPARG*, *NTRK,* and *RAS* mutations for nodules of indeterminate cytology, but they did not recommend the routine use of the Afirma GEC to exclude malignancies because validation studies in the form of long-term outcome data are lacking.[6] The findings of a Taiwanese nationwide survey of thyroid fine-needle aspiration cytology[41] found that the commercially available molecular tests (ThyroSeq, Afirma, RosettaGX, ThyGenX, and ThyraMIR) are rarely used because of their high cost. A chart extraction study from Israel published in 2019[42] examined the use of classification strategies for nodules of indeterminate cytology, including the molecular test Afirma GEC, and found that such molecular markers were not widely available in Israel, but that male sex, immigration status from the Ukraine, and smoking status were three factors that could predict malignancy and help stratify and triage indeterminate nodules. A survey

by the Spanish Society of Endocrinologists and Nutrition in 2018[43] found that molecular testing was used in 8.1% of cases, but it was not clear if these were all nodules of indeterminate cytology, or how widely available molecular testing is. The same survey found that 35.5% of respondents would be willing to undergo surgery for nodules classified TBSRTC III, and 95.8% for nodules classified TBSRTC IV.[43]

Equity

We considered relevant equity issues across different populations defined by the PROGRESS-Plus[44] categories identified during the review process, and we detected no potential health inequities related to the effect of molecular testing for thyroid nodules of indeterminate cytology during scoping.

Thyroid cancer affects more women than men—an estimated three times more; in Ontario in 2018, there were 2,595 cases in women and 746 in men.[11] The rate of increase in thyroid cancer diagnoses is also higher in women than men: the incidence rate in women tripled between 1975 and 2009, and doubled for men in that same time frame. Mortality rates have been stable and similar between the two sexes.[8] The higher observed incidence in females than males is thought to be due in part to differences in hormone levels (including thyroid-stimulating hormone and sex steroids), and to differences in seeking medical attention.[5]

Given that the test is available only as an out-of-pocket expense, and that the test is quite expensive ($2,000 or more), it is very likely that access to these tests is limited to those of higher socioeconomic status; there may also be an inherent bias in whom the test is offered to, based on the judgment of health care providers.

Inequities in representation in study groups for the evaluation of the effect of certain genes may lead to inequities in the application of the tests that rely on those genes. For example, it has been estimated that the prevalence of the *BRAF* gene mutation ranges between 50% and 70% in people from European countries,[12] but the prevalence may be different for people with different racial/ethnic origins. Genetic panels that rely too heavily on the *BRAF* gene to identify malignancies may be less accurate for individuals of different races/ethnicities. This is important, given that people of Asian descent have a higher incidence rate of thyroid cancers.[3]

Expert Consultation

We engaged with experts in the specialty areas of pathology, otolaryngology (head and neck surgery), and genetics to help inform our understanding of aspects of the health technology and our methodologies, and to contextualize the evidence.

PROSPERO Registration

This health technology assessment has been registered in PROSPERO, the international prospective register of systematic reviews (CRD42020214627), available at https://www.crd.york.ac.uk/PROSPERO.

Clinical Evidence

Research Question

What are the clinical validity and clinical utility of molecular testing for people with thyroid nodules of indeterminate cytology?

Methods

Clinical Literature Search

We performed a clinical literature search on September 9, 2020, to retrieve studies published from January 1, 2019, until the search date. We used the Ovid interface in the following databases: MEDLINE, Embase, the Cochrane Central Register of Controlled Trials, the Cochrane Database of Systematic Reviews, the Health Technology Assessment database, and the National Health Service Economic Evaluation Database (NHSEED).

During scoping we learned that Canadian colleagues at the Quebec Institut national d'excellence en santé et en services sociaux (INESSS)[18] were completing a health technology assessment on this topic, and we decided to leverage this work from our pan-Canadian colleagues in completing our review. We based our search strategies on those from INESSS but expanded them to augment return rate. We used controlled vocabulary (e.g., Medical Subject Headings) and relevant keywords designed to capture the population and intervention.

We created database auto-alerts in MEDLINE and Embase and monitored them for the duration of the assessment period. We also performed a targeted grey literature search of health technology assessment agency websites as well as clinical trial and systematic review registries. See Appendix 1 for our literature search strategies, including all search terms.

Eligibility Criteria

STUDIES

Inclusion Criteria

- English-language full-text publications
- Studies published between January 1, 2019, and September 9, 2020
- Systematic reviews, meta-analyses, or health technology assessments

Exclusion Criteria

- Animal and in vitro studies
- Nonsystematic reviews, narrative reviews, abstracts, editorials, letters, case reports, commentaries, or unpublished data

PARTICIPANTS

- Adults (≥ 18 years) with a thyroid nodule of indeterminate cytology:
  - The Bethesda System for Reporting Thyroid Cytopathology (TBSRTC) III: atypical cell of undetermined significance or follicular lesion of undetermined significance (AUS/FLUS)
  - TBSRTC IV: follicular neoplasm/suspicion for a follicular neoplasm (FN/SFN)
  - Other classifications considered reasonably equivalent to TBSRTC III or IV

INTERVENTIONS

- Intervention: a multi-gene panel to support the rule-out of suspected cancerous thyroid nodules prior to surgical removal
  - Included: tests currently available for use by people in Ontario (a commercially available test that involves mailing samples away for testing; a kit with Health Canada approval; or a laboratory-developed test that can be implemented by Ontario laboratories). Known commercially available panels: Afirma (United States; Veracyte gene expression classifier [GEC] or genomic sequencing classifier [GSC]); ThyroSeq (United States; University of Pittsburgh Schools of the Health Sciences, UPMC); ThyGenX, (United States; Interpace Diagnostics Inc., newest version: ThyGeNEXT with ThyraMIR)
  - Excluded: molecular tests as diagnostic or prognostic assessments (i.e., "rule-in" use) and proof-of-concept or gene discovery studies (a commercial brand, Rosetta GX, was excluded because it declared bankruptcy and was taken over by Interpace Diagnostics Inc.,[12] and because we assumed it would not be available or supported in Ontario; it was also primarily a rule-in test)
- Comparator: usual care
  - Included: active surveillance; repeat fine-needle aspiration or thyroidectomy; we considered comparative assessment of the tests identified if data were available
  - Excluded: tests that did not have Health Canada approval (if required) or were currently under research conditions; single gene assessments (e.g., for adding a new gene to an existing gene panel); alternative styles of nodule assessment (e.g., [99mTc]-methyloxyisobutyleisonitirle scintigraphy
- Reference standard: histopathology

OUTCOME MEASURES

- Clinical validity, including all measures of diagnostic accuracy (i.e., determination of whether a nodule is benign or malignant) such as sensitivity, specificity, and predictive values
- Clinical utility, including measures of patient outcomes (such as disease progression) and measures of health care utilization (such as surgical resection rates)

Outcomes related to patient preferences were not included in this systematic review of the clinical evidence; they are explored in the Quantitative Evidence section of the Preferences and Values chapter, later in this report.

Literature Screening

A single reviewer conducted an initial screening of titles and abstracts using Covidence[45] and then obtained the full texts of studies that appeared eligible for review according to the inclusion criteria. A single reviewer then examined the full-text articles and selected studies eligible for inclusion. A single reviewer also examined reference lists and consulted content experts for any additional relevant studies not identified through the search.

Data Extraction

We extracted relevant data on study characteristics and risk-of-bias items using a data form to collect information on the following:

- Source (e.g., citation information, study type)
- Methods (e.g., study design, study duration and years, participant allocation, allocation sequence concealment, blinding, reporting of missing data, reporting of outcomes, whether the study compared two or more groups)
- Outcomes (e.g., outcomes measured, number of participants for each outcome, number of participants missing for each outcome, outcome definition and source of information, unit of measurement, upper and lower limits [for scales], time points at which the outcomes were assessed)

We contacted study authors to provide clarification as needed.

Statistical Analysis

We did not conduct meta-analysis because the available evidence was limited; all findings are reported narratively.

Critical Appraisal of Evidence

We sought reviews that conducted their own critical appraisal of the evidence, and we have reported the risk-of-bias assessments as conducted by those reviews (Appendix 2). We assessed risk of bias in systematic reviews using the Risk of Bias in Systematic Review (ROBIS)[46] tool. We assessed risk of bias in primary studies using the Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2)[47] assessment for studies that evaluated clinical validity (diagnostic accuracy) and the Risk of Bias tool for Non-randomized Studies (RoBANS)[48] for studies that evaluated clinical utility.

We evaluated the quality of the body of evidence for each outcome according to the *Grading of Recommendations Assessment, Development, and Evaluation* (GRADE) *Handbook*.[49] The body of evidence was assessed based on the following considerations: risk of bias, inconsistency, indirectness, imprecision, and publication bias. The overall rating reflects our certainty in the evidence.

Results

Clinical Literature Search

The database search of the clinical literature yielded 430 citations published from database inception until September 9, 2020. We identified two additional eligible studies from other sources (one of these was published in French, but because it was a publication from a Canadian health technology assessment agency, we chose to include it). In total, we identified seven systematic reviews that met our inclusion criteria. Figure 2 presents the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) flow diagram for the clinical literature search.

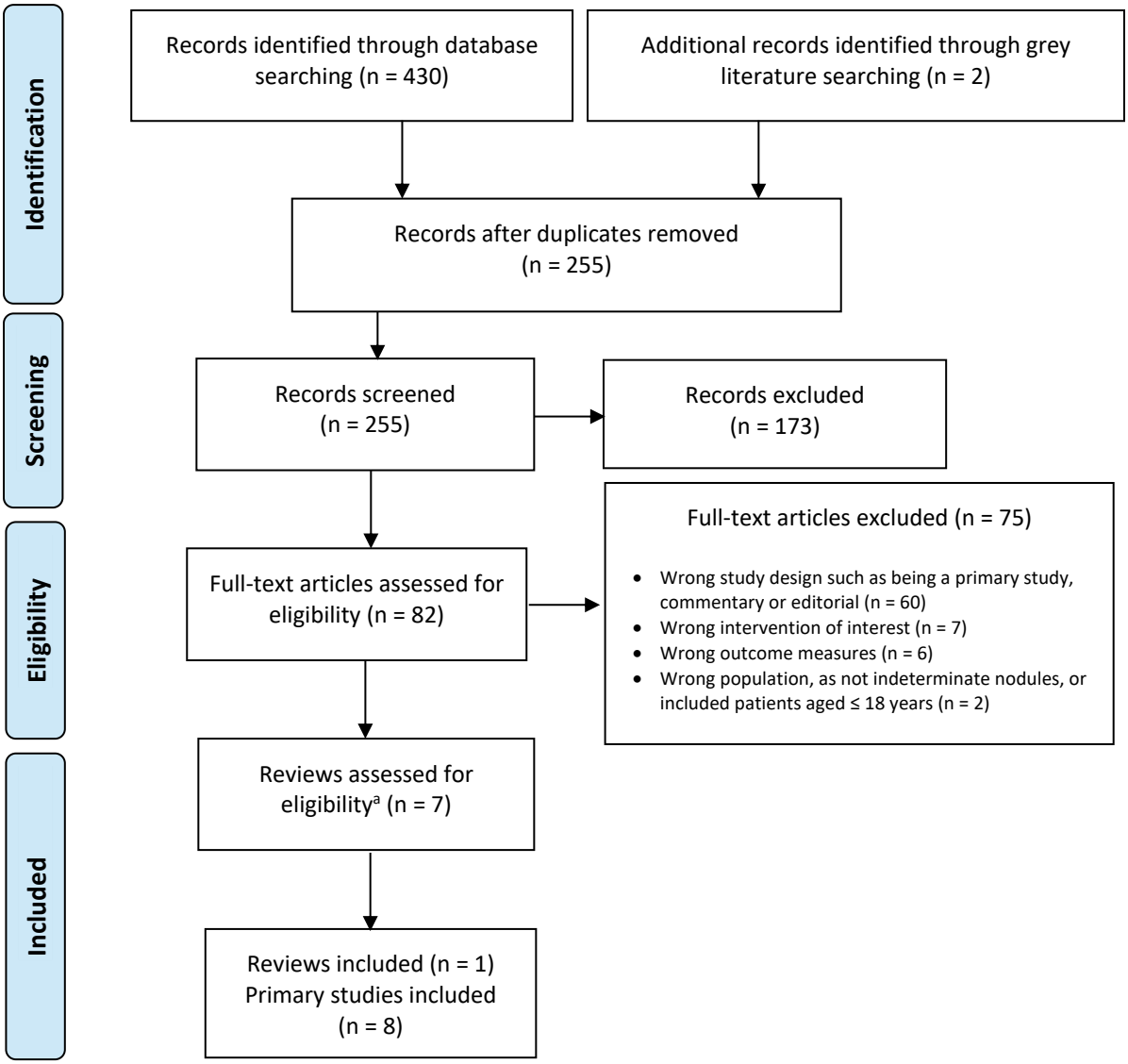

Figure 2: PRISMA Flow Diagram—Clinical Search Strategy

Abbreviation: PRISMA, Preferred Reporting Items for Systematic Reviews and Meta-analyses.

[a] After full-text review and Risk of Bias in Systematic reviews (ROBIS) assessment, we selected one systematic review because it was the most contextually relevant, the most recent, and of the best quality.

Source: Adapted from Moher et al.[50]

Characteristics of Included Studies

SYSTEMATIC REVIEWS

We identified seven systematic reviews that addressed our research question. We considered all reviews for their relevance, age, and quality as assessed by ROBIS.[46] The report published by INESSS[18] was the most complete, encompassing all relevant evidence up to date (including the most recently published available evidence); it was also contextually relevant, having been developed from the perspective of another Canadian province. In this health technology assessment, we have focused on the findings from the INESSS report; a summary of the other reviews is available in Appendix 3.

PRIMARY STUDIES

The INESSS report identified 46 primary studies[18] that included several commercially available brands of molecular thyroid tests and their various versions over time. However, INESSS limited its conclusions to the two molecular tests that had published clinical validation studies at the time of its review. This health technology assessment also focuses on the findings for the most recent versions of these tests: Afirma GSC and ThyroSeq v3. A summary of the eight primary studies that evaluated these two tests and were included in the INESSS report is presented in Table 2, two studies evaluated clinical validity and clinical utility outcomes.

Table 2: Characteristics of Primary Studies[a]

Author, year	Test type	Study design	% Female	Mean age, y	Sample size (nodules), n	Average number of nodules per person	TBSRTC III, n (%)	TBSRTC IV, n (%)	Median nodule size (range), cm[b]
Angell et al, 2019[51]	Afirma GSC	Retrospective	84.6	57 (median)	114	1.06	67 (59)	47 (41)	2.0 (1.0–6.9)
Chen et al, 2020[33]	ThyroSeq v3	Retrospective	26	54	48	NR	20 (42)	28 (58)	2.5 (1.1–4.2)
Endo et al, 2019[52]	Afirma GSC	Retrospective	75	54.8[c]	164	1.08	124 (76)	40 (24)	2.3 (NR)[c]
Harrell et al, 2019[53]	Afirma GSC	Retrospective	NR	NR	139	NR	NR	NR	NR
Patel et al, 2018[54]	Afirma GSC	Retrospective, multicentre	77.6	51.7	190	NR[d]	114 (60)	76 (40)	2.6 (1.0–9.1)
San Martin et al, 2020[55]	Afirma GSC	Retrospective	75[e]	56.1[e]	121	1.04	76 (63)	45 (37)	2.0 (1.4–3.2)
Steward et al, 2019[56]	ThyroSeq v3	Prospective, multicentre	80	53 (median)	257	1.18	154 (62)	93 (38)	2.4 (0.5–8)
Wei et al, 2019[57]	Afirma GSC	Retrospective	73	57.7[c]	78	1.0	66 (85)	12 (15)	NR

Abbreviations: GEC, gene expression classifier; GSC, gene sequence classifier; INESSS, Institut national d'excellence en santé et en services sociaux; NR, not reported; TBSRTC, The Bethesda System for Reporting Thyroid Cytology.

[a] Information was sourced from the INESSS report by the authors of this health technology assessment,[18] referring to primary studies as needed.

[b] From ultrasound.

[c] Calculated based on information from the primary study.

[d] 1.2 nodules per person for the initial cohort that received the GEC test; this study assessed a subgroup of those who had also received the GSC.

[e] Reported for 116 of the 121 nodules included in the study.

Risk of Bias in the Included Studies

INESSS[18] assessed the quality of the studies that evaluated clinical validity (diagnostic accuracy) using QUADAS-2 (adapted from another review by Duh et al[58]). We adapted these findings for our assessment and supplemented them with information from the primary studies as needed.

Our risk-of-bias assessment (Appendix 2) found a high risk of bias in the two clinical validity studies[54,56] (based on QUADAS-2[59]) and the six clinical utility studies[33,51-53,55,57] (based on RoBANS[48]).

Clinical Validity

Measures of diagnostic accuracy (sensitivity and specificity), using data from the INESSS report[18] (and from primary studies as required), are shown in Table 3.

Table 3: Molecular Testing for Thyroid Nodules of Indeterminate Cytology—Diagnostic Accuracy[a]

	Sensitivity (95% CI), %			Specificity (95% CI), %		
Molecular test	**TBSRTC III**	**TBSRTC IV**	**Combined**	**TBSRTC III**	**TBSRTC IV**	**Combined**
Afirma GSC (1 study[54])	93 (76–99)	88 (64–99)	91 (79–98)	71 (60–80)	64 (51–76)	68 (60–76)
ThyroSeq v3 (1 study[56])	91 (77–97)	97 (85–100)	94 (86–98)	85 (77–90)	75 (63–84)	82 (75–87)

Abbreviations: CI, confidence interval; GSC, gene sequencing classifier; TBSRTC, The Bethesda System for Reporting Thyroid Cytopathology.

[a] Classification of a thyroid nodule as benign (TBSRTC II), or malignant (TBSRTC VI).

For clinical validity (diagnostic accuracy), our GRADE certainty in the body of evidence was Low for the two molecular tests assessed; the evidence was downgraded because the studies had serious limitations related to risk of bias—specifically the selection of participants and publication bias (Appendix 2, Tables A2 and A4).

We calculated predictive values for the Afirma GSC and ThyroSeq v3 using a range of prevalence rates based on those reported in the original primary studies (Table 4). We have also reported a measure commonly used by the manufacturers: the rate at which the molecular tests returned a null (benign) finding, known as the "benign cell rate." This rate would be expected to be a perfect inverse of the prevalence rate if the test had perfect accuracy.

Table 4: Molecular Testing for Thyroid Nodules of Indeterminate Cytology—Benign Cell Rates and Predictive Values

Molecular test	Benign cell rate[a]	Prevalence of malignant nodules[b]	Predictive values using reported prevalence rates[b]		Predictive values using estimated prevalence rates[c]					
					20%		25%		35%	
			PPV	NPV	PPV	NPV	PPV	NPV	PPV	NPV
Nodules of indeterminate cytology[d]										
Afirma GSC 1 study[54]	54%	24%	47%	96%	39%	97%	49%	96%	61%	93%
ThyroSeq v3 1 study[56]	61%	28%	66%	97%	56%	98%	63%	98%	73%	96%
TBSRTC III										
Afirma GSC 1 study[54]	55%	25%	51%	97%	44%	98%	52%	97%	63%	95%
ThyroSeq v3 1 study[56]	68%	23%	64%	97%	60%	98%	67%	97%	76%	95%
TBSRTC IV										
Afirma GSC 1 study[54]	53%	22%	42%	95%	38%	96%	45%	94%	57%	91%
ThyroSeq v3 1 study[56]	49%	35%	68%	98%	49%	99%	56%	99%	68%	98%

Abbreviations: GSC, GSC, gene sequencing classifier; NPV, negative predictive value; PPV, positive predictive value; TBSRTC, The Bethesda System for Reporting Thyroid Cytology.

[a] Range of benign molecular test findings.

[b] Calculations based on data from the primary studies.

[c] Values calculated based on prevalence rates deemed reasonable for the Ontario context.

[d] Combined TBSRTC III and TBSRTC IV cohorts.

Clinical Utility

The only measure of clinical utility evaluated in the INESSS report[18] was rate of resection (surgery) after a benign molecular test result. Resection rates for nodules with benign findings appeared to be lower than for all findings of indeterminate cytology (Table 5; observed ranges 27% to 44% and 6% to 7%, respectively). In comparison, the estimated resection rate without molecular testing is approximately 55% for nodules of indeterminate cytology (although rates vary by centre),[16] in alignment with data from hospital administrative databases in Ontario but lower than the estimates of experts, who thought resection rates were above 70% (Antoine Eskander, email communication, March 28, 2021).

We were also able to compare resection rates for nodules classified as benign based on molecular testing and nodules confirmed to be benign using cytology (i.e., TBSRTC II). The resection rate for benign nodules based on molecular testing (6% to 7%; Table 5) was slightly lower than the resection rate for benign nodules based on cytology (with no molecular testing; approximately 10.4%).[16] (Estimates of absolute numbers for Ontario based on the above resection rates have been developed for the economic models used later in this report). We identified no studies that examined this measure directly for the most recent versions of the molecular tests of interest, but studies using the older version of the Afirma test (GEC) did explore this outcome and found similar resection rates.[18]

Table 5: Resection Rate for Thyroid Nodules of Indeterminate Cytology After Molecular Testing

Molecular test	Nodules found to be benign	Resection rate[a]			
		TBSRTC III	TBSRTC IV	Nodules of indeterminate cytology (TBSRTC III and IV)	Nodules found to be benign based on molecular testing
Afirma GSC 5 studies [51-53,55,57]	68% (419 of 616)	NR	NR	27% (168 of 616)	6% (26 of 419)
ThyroSeq v3 1 study[33]	58% (28 of 48)	NR	NR	44% (21 of 48)	7% (2 of 28)

Abbreviations: CI, confidence interval; GSC, gene sequencing classifier; NR, not reported; TBSRTC, The Bethesda System for Reporting Thyroid Cytopathology.

[a] Among nodules assessed using other molecular tests.

Resection rates were not reported separately for nodules classified as TBSRTC III and IV. Among the nodules found to be benign based on molecular testing with the Afirma GSC but then resected, 19% (5 of 26) were found to be malignant or noninvasive follicular thyroid neoplasms with papillary nuclear characteristics after histopathological testing. The two nodules that were classified as benign using the ThyroSeq v3 test but then resected were confirmed to be benign with histopathology.

For clinical utility (resection rate), our GRADE certainty in the body of evidence was Very low for both of the molecular tests assessed; the evidence was downgraded because the included studies were observational and had serious limitations related to risk of bias—specifically the selection of participants and publication bias (Appendix 2, Tables A3 and A4).

Ongoing Studies

We are aware of one ongoing study assessing a molecular classifier for thyroid nodules, known as the ThyroPred-1. This study is registered at ClinicalTrials.gov (NCT003392402).[60]

Discussion

Molecular testing for nodules of indeterminate cytology—specifically TBSRTC III and IV nodules—may be useful in reducing rates of unnecessary surgery. The two tests included in this health technology assessment (Afirma GSC and ThyroSeq v3) had sensitivities of 91% to 94% and specificities of 68% to 82%, respectively, for the detection of malignancy. With molecular testing (either test), resection rates were reduced for nodules classified as benign compared to nodules classified as malignant, and comparable to resection rates for nodules classified as benign (e.g., TBSRTC II) based on cytology.

Nevertheless, contextualizing the evidence to clinical practice in Ontario requires us to consider the fact that surgery is not conducted only to manage or rule out malignancy. People may undergo surgery for a number of reasons, including anxiety; respiratory and/or digestive complications resulting from a growth; or symptoms from conditions such as uncontrolled hyperthyroidism. As a result, molecular testing may not lead to reduced resection rates for nodules of indeterminate cytology that are comparable to those reported in the included studies. The INESSS recommendation[18] stipulates that molecular testing of thyroid nodules should be considered only in circumstances in which it is likely to affect decisions about surgery.

The molecular tests identified in this health technology assessment use TBSRTC to identify patients with indeterminate cytology; if molecular testing were implemented in Ontario, it would be important to take into account the fact that TBSRTC is not used consistently across Ontario to report cytology (Katie O'Reilly, email communication, April 14, 2021). Standardization of reporting may be a key first step in implementation, based on the protocols of the studies that evaluated the molecular tests we included in this review. As well, because there is nothing proprietary about the commercially available tests, an Ontario-based, laboratory-developed test could prove just as effective at a lower cost (Antoine Eskander, email communication, March 28, 2021; Aaron Pollett, email communication, August 16, 2021).

Strengths and Limitations

The body of evidence and our methodological approach had several limitations.

First, a comparison of the raw data may suggest that there could be differences between the Afirma GSC and ThyroSeq V3 tests in terms of diagnostic accuracy and resection rates. However, the studies that evaluated these tests had limitations in their designs and applied slightly different methodologies and inclusion criteria, so we considered it inappropriate to draw conclusions from observed comparisons of their findings.

As well, the included primary studies were limited in the populations they studied. They were focused largely on general adult populations. Other specific groups develop thyroid nodules, such as children and pregnant people, and although treatment may generally be similar for these groups, molecular testing has not yet been validated in these populations.[61] As well, the included primary studies did not report the racial/ethnic backgrounds of their study participants, so it is uncertain whether the study populations were representative of those expected to experience thyroid nodules of indeterminate cytology in Ontario. People of Asian descent have higher incidence rates of thyroid cancer,[3]

but historically, people of European descent have been overrepresented in genetic studies, leading to prevalence or malignancy estimates that may be inaccurate for people of different racial/ethnic origins. For example, genetic panels that rely too much on the *BRAF* gene to identify potential malignancies may be inequitable in their effectiveness for people of different racial/ethnic groups.

The study population of one of the included studies, Chen et al,[33] had baseline characteristics that were different from the other studies: only 26% were female, and the population included more people with TBSRTC IV nodules than TBSRTC III, which was atypical. This study was the only one to evaluate ThyroSeq v3 resection rates; we accounted for these limitations of the body of evidence in our GRADE assessment. During the publication phase of our report, we identified two studies[62,63] that were published after the INESSS report and therefore not included in their review; these studies had findings similar to our report with respect to accuracy and reduced resection rates among people who had received molecular testing. As well, in our calculations for predictive values we evaluated a range of prevalence estimates based on those observed in the primary studies. These estimates may not be representative of the true prevalence rates in Ontario; however, they are likely to be conservative for our consideration of molecular testing as rule-out tests, and the findings for negative predictive value were already high.

Our methodological approach was to leverage the findings of published systematic reviews, so we were limited by the methods and findings of the review we included. Still, our approach of conducting a systematic search to identify the most relevant and high-quality review ensured that we leveraged the best available evidence. Furthermore, the review by INESSS[18] limited its assessment of clinical validity to the best available evidence, including only studies in which all nodules evaluated using molecular testing were also resected for confirmatory diagnostic assessment. A large body of evidence in this field did not meet these criteria. Instead, the evidence focused on people with benign findings from molecular testing that were typically managed using active surveillance; samples did not undergo confirmatory histopathology diagnostics, affecting the findings of these validation studies. This methodological approach may have skewed resection rates toward higher-risk cases, meaning that findings were not necessarily representative of the real-world effectiveness of the tests. The best method for evaluating the performance of a test is by validating its findings against a reference standard (in this case, histopathological testing). As such, although INESSS omitted some studies that were included in other reviews, we agreed with their decision to focus on studies with the best available evidence and to require confirmatory diagnostic testing with a reference standard.

The INESSS report[18] also limited its review to rule-out tests because the authors felt that molecular testing was not well established for prognostication and therapy decisions as this time. Their decision aligned with guidance in the field, such as that from the American Thyroid Association, which suggests that a rule-in test would require a positive predictive value similar to that for malignant cytological diagnosis, which is 98.6%.[9] We confirmed with Ontario clinical experts that this focus on rule-out tests was a reasonable approach for the state of the evidence at present.

We focused on research findings for the most recent versions of the two tests that are currently available in Ontario. We did not explore the body of evidence for the initial iterations of these tests. For example, more studies have evaluated the Afirma GEC than the more recently developed GSC. The INESSS report[18] did explore the evidence for the previous versions. We felt that the evidence was sufficient to show that the newer versions were comparable to their predecessors, or improvements on them.

Finally, this health technology assessment evaluated the use of molecular testing for thyroid nodules in a rule-out capacity to reduce the risk of unnecessary surgery; however, other technological advances are also aimed at improving the overall diagnostic process for thyroid nodules. For example, new gene candidates are continually being evaluated, and some are being targeted for use as part of rule-in tests and preoperative diagnostic tools.[64-68] As noted above, it may also be of value to consider developing a test based in Ontario laboratories that uses established knowledge in this field. In addition to specific genes, the use of advanced genetic evaluation techniques such as droplet digital polymerase chain reaction and a more epigenetic focus in testing should also be considered.[69,70] As well, advances in collection and smear techniques for fine-needle aspiration can minimize the need for repeat sample collection.[71,72] Furthermore, advances in diagnostic tools may eventually replace fine-needle aspiration with tests such as liquid biopsy[73,74]; including combinations of circulating micro RNA and sonographic thyroid imaging reporting and data systems (TI-RADS)[75,76]; improved ultrasound risk-stratification techniques[77]; mass spectrometry[78]; machine-learning texture analyses[79]; and constantly improving methods of cytological examination.[80]

Conclusions

Molecular testing for thyroid nodules of indeterminate cytology has a sensitivity of 91% to 94% and a specificity of 68% to 82% for the detection of malignancy. As well, lower rates of surgical resections were reported in nodules of indeterminate cytology compared to usual (molecular testing vs. no molecular testing), but the evidence is very uncertain.

Economic Evidence

Research Question

What is the cost-effectiveness of molecular testing compared to usual care for people with thyroid nodules of indeterminate cytology?

Methods

Economic Literature Search

We performed an economic literature search on September 10, 2020, to retrieve studies published from database inception until the search date. To retrieve relevant studies, we developed a search using the clinical search strategy with an economic and costing filter applied. Although the clinical search used a date limit of 2019, our preference was to review all economic studies, regardless of the date of publication.

We created database auto-alerts in MEDLINE and Embase and monitored them for the duration of the assessment period. We also performed a targeted grey literature search of health technology assessment agency websites, clinical trial and systematic review registries, and the Tufts Cost-Effectiveness Analysis Registry. See the Clinical Literature Search section, above, for further details on methods used. See Appendix 1 for our literature search strategies, including all search terms.

Eligibility Criteria

STUDIES

Inclusion Criteria

- English-language full-text publications
- Studies published from database inception until September 10, 2020
- Cost–benefit analyses, cost-effectiveness analyses, cost-minimization analyses, or cost–utility analyses

Exclusion Criteria

- Abstracts, case reports, editorials, commentaries, reviews, letters, unpublished studies
- Cost analyses

PARTICIPANTS

- Adults (≥ 18 years) with a thyroid nodule of indeterminate cytology

INTERVENTIONS

- Intervention: a multi-gene panel to support the rule-out of suspected cancerous thyroid nodules prior to surgical removal
  - Included: tests currently available for use by people in Ontario (a commercially available test that involves mailing samples away for testing; a kit with Health Canada approval; or a laboratory-developed test that can be implemented by Ontario laboratories). Known commercially available panels: Afirma (United States; Veracyte gene expression classifier [GEC] or genomic sequencing classifier [GSC]); ThyroSeq (United States; University of Pittsburgh Schools of the Health Sciences, UPMC); ThyGenX, (United States; Interpace Diagnostics Inc., newest version: ThyGeNEXT with ThyraMIR)
  - Excluded: molecular tests as diagnostic or prognostic assessments (i.e., "rule-in" use) and proof-of-concept or gene discovery studies (a commercial brand, Rosetta GX, was excluded because it declared bankruptcy and was taken over by Interpace Diagnostics Inc.,[12] and because we assumed it would not be available or supported in Ontario; it was also primarily a rule-in test); studies that evaluated preliminary/earlier versions of the molecular tests
- Comparator: usual care
  - Included: active surveillance; repeat fine-needle aspiration or thyroidectomy
  - Excluded: studies in which the usual care strategy no longer reflected the current management of thyroid nodules (e.g., clinical management based on the 2009 American Thyroid Association guidelines,[81] rather than the 2015 revision[9])

OUTCOME MEASURES

- Costs
- Health outcomes (e.g., quality-adjusted life-years [QALYs], life-years, rate of unnecessary surgery, probability of predicting a correct diagnosis)
- Incremental costs
- Incremental effectiveness
- Incremental cost-effectiveness ratios (ICERs)

Literature Screening

A single reviewer conducted an initial screening of titles and abstracts using Covidence[45] and then obtained the full texts of studies that appeared eligible for review according to the inclusion criteria. The same reviewer then examined the full-text articles and selected studies eligible for inclusion.

Data Extraction

We extracted relevant data on study characteristics and outcomes to collect information about the following:

- Source (e.g., citation information, study type)
- Methods (e.g., study design, analytic technique, perspective, time horizon, population, intervention[s], comparator[s])
- Outcomes (e.g., health outcomes, costs, incremental cost-effectiveness ratios)

Study Applicability

We determined the usefulness of each identified study for decision-making by applying a modified quality appraisal checklist for economic evaluations originally developed by the National Institute for Health and Care Excellence (NICE) in the United Kingdom to inform the development of NICE's clinical guidelines.[82] We modified the wording of the questions to remove references to guidelines and to make it specific to Ontario.

Results

Economic Literature Search

The database search of the economic literature search yielded 179 citations published from database inception until September 10, 2020. We identified six additional studies from other sources (one of these was published in French, but because it was a publication from a Canadian health technology assessment agency, we chose to include it). In total, we identified two cost-effectiveness studies that met our inclusion criteria. Figure 3 presents the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) flow diagram for the economic literature search.

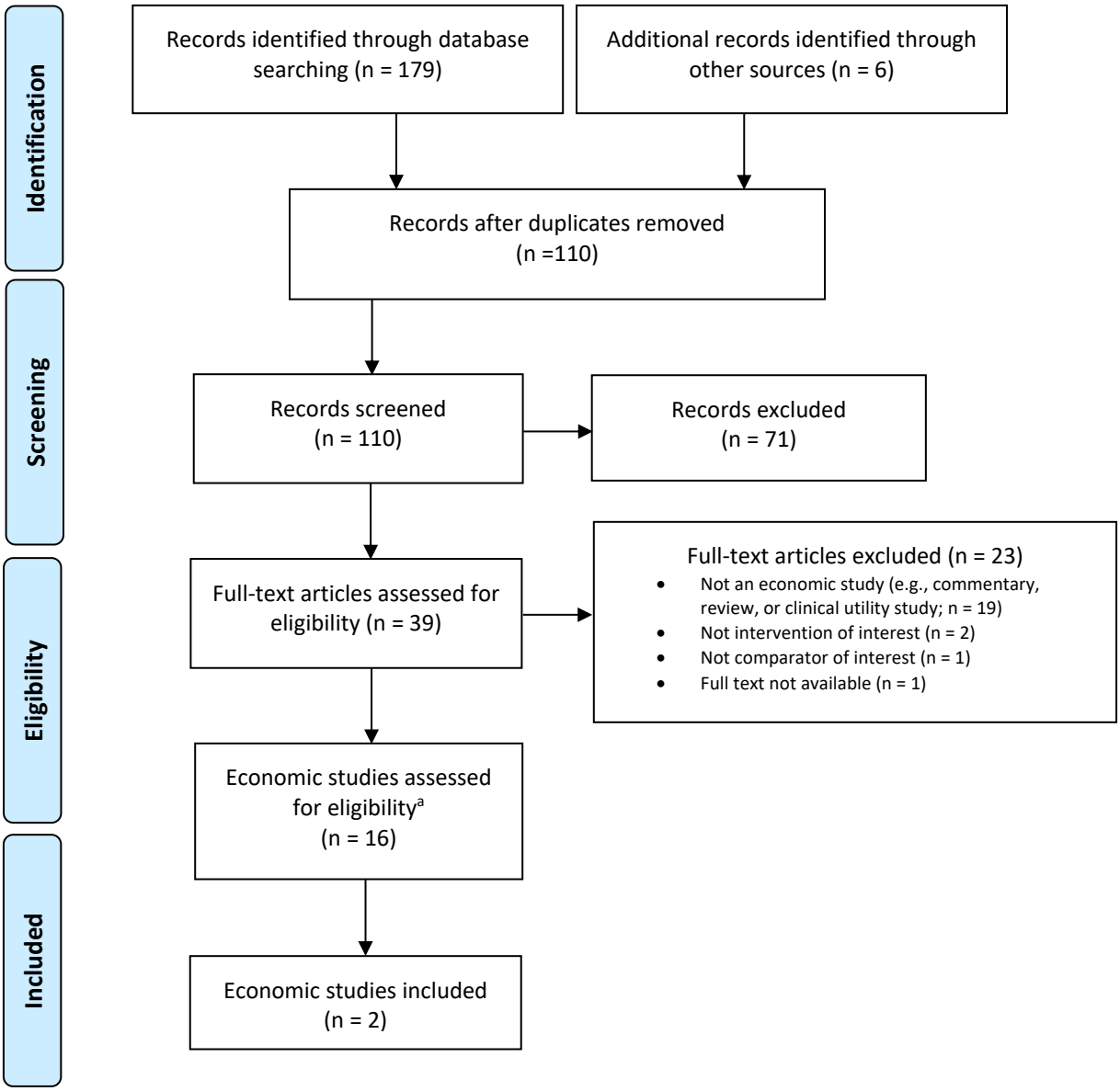

Figure 3: PRISMA Flow Diagram—Economic Search Strategy

Abbreviation: PRISMA, Preferred Reporting Items for Systematic Reviews and Meta-analyses.

[a] After full-text review, 2 were selected as the most contextually relevant to the current decision problem.

Source: Adapted from Moher et al, 2009.[50]

Overview of Included Economic Studies

Although we initially identified 16 economic studies on molecular testing of thyroid nodules, we narrowed the selection to only those that were most relevant to the current decision problem (see Appendix 4 for a list of excluded studies). We excluded studies in which molecular testing was used as a rule-in test,[83-86] because according to clinical experts, rule-in tests are still too early in their development to be adopted into clinical practice. We also excluded studies that evaluated preliminary or earlier versions of the available molecular tests (e.g., ThyroSeq v2, Afirma GEC), whose test performance may have differed from the current versions (e.g., ThyroSeq v3, Afirma GSC).[83-94] Finally, we excluded studies in which usual care no longer reflected the current management of thyroid nodules (e.g., studies that followed the 2009 American Thyroid Association guidelines,[81] rather than the 2015 revision[9]). Two studies met the inclusion criteria; the results of these studies are summarized in Table 6.

Nicholson et al[95] compared the costs and effectiveness of Afirma GSC, ThyroSeq v3, and diagnostic lobectomy from a US payer perspective over a 20-year time horizon. The study included the costs of lobectomy, lobectomy-associated complications (hematoma, hypothyroidism, and vocal cord dysfunction), surveillance (follow-up visits and thyroid ultrasound), molecular testing, and a second fine-needle aspiration biopsy related to molecular testing. They found that Afirma GSC and ThyroSeq v3 led to a higher probability of predicting a correct diagnosis compared to diagnostic lobectomy (63.7% and 73.2% vs. 25%), but the two molecular tests were associated with higher costs ($11,385 USD and $10,451 USD per patient vs. $9,602 USD per patient). The cost per correct diagnosis was $14,277 USD for ThyroSeq v3, $17,873 USD for Afirma GSC, and $38,408 USD for diagnostic lobectomy. A probabilistic sensitivity analysis showed that ThyroSeq v3 was the preferred management strategy in 68.5% of cases.

We also identified an economic analysis conducted by the Quebec Institut national d'excellence en santé et services sociaux (INESSS) as part of a health technology assessment on molecular testing of thyroid nodules.[18] The analysis compared the costs and QALYs associated with Afirma GSC, ThyroSeq v3, and diagnostic lobectomy from a Quebec public payer perspective over a time horizon of 5 years. The analysis included the costs of lobectomy, lobectomy-associated complications (hypothyroidism), surveillance, and molecular testing. The analysis found that the ThyroSeq v3 strategy was slightly less costly than the diagnostic lobectomy strategy ($5,972 vs. $6,009 per patient over 5 years), and it was more effective (4.381 vs. 4.359 QALYs over 5 years; i.e., ThyroSeq v3 dominated diagnostic lobectomy). The Afirma GSC strategy was slightly more costly and more effective than diagnostic lobectomy ($6,287 per patient and 4.371 QALYs over 5 years), resulting in an ICER of $22,667 per QALY. At a willingness-to-pay value of $50,000 per QALY, ThyroSeq v3 was likely to be cost-effective (69% probability), but the cost-effectiveness of Afirma GSC was uncertain (52% probability).

Table 6: Results of Economic Literature Review—Summary

Author, year, country	Analytic technique, study design, perspective, time horizon	Population	Intervention and comparator	Results		
				Health outcomes	Costs	Cost-effectiveness
Nicholson et al, 2019,[95] United States	Cost-effectiveness analysis Decision tree and Markov model Health care payer perspective (Medicare) Time horizon: 20 y Discount rate: 3% (cost only)	Adults with thyroid nodules of indeterminate cytology (TBSRTC III and IV) Age: 40 y Female: 100% Nodule size: 2 cm	Afirma GSC ThyroSeq v3 DL[a]	*Probability of predicting a correct diagnosis*[b] Afirma GSC: 0.637 ThyroSeq v3: 0.732 DL: 0.250 *Incremental probability of predicting a correct diagnosis vs. DL (calculated)* Afirma GSC: 0.387 ThyroSeq v3: 0.482	*Total cost per patient, 2018 USD* Afirma GSC: $11,385 ThyroSeq v3: $10,451 DL: $9,602 *Incremental cost vs. DL (calculated)* Afirma GSC: $1,783 ThyroSeq v3: $849 (Unit cost of molecular test: $3,600; unit cost of DL: $9,520)	Compared to DL, both molecular tests led to a higher probability of predicting a correct diagnosis but were more costly. Cost per correct diagnosis: $14,277 for ThyroSeq v3, $17,873 for Afirma GSC, and $38,408 for DL *Probabilistic analysis* ThyroSeq v3 was the preferred strategy in 68.5% of cases (vs. 25% for Afirma GSC and 6.5% for DL) *Sensitivity analyses* Result was most sensitive to variations in molecular test sensitivity, cancer prevalence, and probability of test failure
INESSS, 2021,[18] Quebec, Canada	Cost–utility analysis Decision tree and Markov model Health care payer perspective Time horizon: 5 y Discount rate: NR	Adults with thyroid nodules of indeterminate cytology (TBSRTC III and IV) Age: NR Female: NR Nodule size: NR	Afirma GSC ThyroSeq v3 DL[c]	*Total QALYs per patient* Afirma GSC: 4.371 ThyroSeq v3: 4.381 DL: 4.359 *Incremental QALYs vs. DL* Afirma GSC: 0.012 ThyroSeq v3: 0.022	*Total cost per patient, 2020 CAD* Afirma GSC: $6,287 ThyroSeq v3: $5,972 DL: $6,009 *Incremental cost vs. DL* Afirma GSC: $272 ThyroSeq v3: −$37 (Unit cost of molecular test: $5,385; unit cost of DL: $9,842)	ThyroSeq v3 was slightly less costly (savings of $37/patient) and slightly more effective (0.022 QALY gained) than DL Afirma GSC was more costly ($272/patient) and slightly more effective (0.012 QALY gained) than DL, with an ICER of $22,667/QALY *Probabilistic analysis* At a willingness-to-pay value of $50,000/QALY, ThyroSeq v3 was likely to be cost-effective (69% probability); the cost-effectiveness of Afirma GSC was uncertain (52% probability)

Abbreviations: DL, diagnostic lobectomy; GSC, gene sequencing classifier; ICER, incremental cost-effectiveness ratio; INESSS, Institut national d'excellence en santé et services sociaux; NR, not reported; QALY, quality-adjusted life-year; TBSRTC, The Bethesda System for Reporting Thyroid Cytopathology.

[a] All patients in this treatment arm received diagnostic lobectomy.

[b] The probability of predicting a correct diagnosis is defined as choosing diagnostic lobectomy for a histologically malignant nodule or choosing surveillance for a histologically benign nodule.

[c] Although the comparator strategy was diagnostic lobectomy, only 55% of patients who received this strategy were modelled to undergo surgery; the remaining patients were modelled to receive surveillance.

Applicability of the Included Studies

Appendix 5 provides the results of the quality appraisal checklist for economic evaluations applied to the included studies. The INESSS analysis was deemed partially applicable to our research question (same population and intervention, but different resource use and costs).

Discussion

The economic evidence for the more recent versions of the molecular tests (ThyroSeq v3 and Afirma GSC) is still very limited: only two studies met our inclusion criteria. Both studies compared molecular testing with diagnostic lobectomy as the usual care strategy, and both studies used the same values for key model parameters, such as cancer prevalence (25%), molecular test sensitivity and specificity, and probability of test failure (10%). However, the two studies were conducted in different settings (one in the United States and one in Quebec, Canada), and the modelling approach was different in several ways. Nicholson et al[95] used a 20-year time horizon, and they assumed that all positive and inconclusive test results would lead to diagnostic lobectomy and all negative test results would lead to surveillance. The INESSS report[18] used a shorter (5-year) time horizon and modelled a more realistic clinical pathway with different surgical resection rates following a positive or negative test result: the authors assumed that 77% of patients with a positive result would undergo surgery and 13% of patients with a negative result would undergo surgery. The two studies also modelled their comparator strategy (diagnostic lobectomy) differently. In the study by Nicholson et al,[95] all patients in the diagnostic lobectomy arm received surgery. In the INESSS report,[18] only 55% of patients in the diagnostic lobectomy arm received surgery; the remaining patients received surveillance. Finally, the two studies measured effectiveness using different outcomes. Nicholson et al[95] estimated the probability of predicting a correct diagnosis (defined as choosing diagnostic lobectomy for a histologically malignant nodule or choosing surveillance for a histologically benign nodule), whereas the INESSS report[18] estimated the QALYs associated with each strategy.

Although both studies were generally well conducted, their results may not be generalizable to Ontario. The study by Nicholson et al[95] was conducted in the United States, so the results cannot be applied to an Ontario context. Their analysis also calculated the cost per correct diagnosis associated with each strategy and deemed molecular testing to be more cost-effective because the cost per correct diagnosis was lower than that of diagnostic lobectomy. However, their findings should have been interpreted using ICERs instead of absolute cost per correct diagnosis. Also, because there is no commonly accepted willingness-to-pay value for incremental cost per correct diagnosis, it is difficult to determine whether molecular testing was cost-effective based on these findings. Although the INESSS analysis[18] was conducted in a Canadian setting, some key input parameters (such as test and surgery costs) were different for Quebec and Ontario. The authors reported that in Quebec, the cost of molecular testing was $5,385 and the cost of diagnostic lobectomy was $9,842. However, in Ontario, the cost of a diagnostic lobectomy in 2020 was only about $5,412, according to the Ontario Case Costing database.[96]

Excluded Studies

The economic analyses that evaluated older versions of the molecular tests (i.e., Afirma GEC, ThyroSeq v2) had mixed results: some studies found molecular testing to be cost-saving[86,89,90,94] or cost-effective,[84] but some studies found the opposite.[83,87,88,91-93] Such variation in results may have been caused by differences in analysis approach, test cost, surgery cost, or the prevalence of malignancy. In studies that found molecular testing to be cost-saving, there was usually a large cost difference between molecular testing and surgery. For example, in an analysis by Yip et al,[86] the cost of the molecular test was $650 (2010 USD), and the cost of diagnostic lobectomy was $7,301 ($6,549 for

hospitalization and $752 for physician fees). In another study by Rivas et al,[89] the cost of the molecular test was $4,056 (USD, costing year not reported) and the cost of diagnostic lobectomy was $20,200. Another study[84] found molecular testing to be cost-effective based on incremental cost per unnecessary surgery avoided, but it is difficult to judge whether a technology is cost-effective based on such findings because there is no commonly used willingness-to-pay value for this outcome.

Conclusions

We found two economic analyses that compared molecular testing with diagnostic lobectomy in patients with thyroid nodules of indeterminate cytology. One study suggested that compared to diagnostic lobectomy, molecular testing could lead to a higher probability of predicting a correct diagnosis but at a higher cost. Another study found that molecular testing was probably cost-effective, but with some uncertainty. However, because of differences in resource use and costs between settings, these findings were not generalizable to an Ontario context.

Primary Economic Evaluation

We identified two cost-effectiveness analyses that compared molecular testing with usual care in people with thyroid nodules of indeterminate cytology. One of the analyses was conducted by the Quebec Institut national d'excellence en santé et en services sociaux (INESSS)[18] in a Canadian health care setting. However, some key input parameters were significantly different for Quebec and Ontario (e.g., the costs of molecular testing and diagnostic lobectomy were both lower in Ontario), so the results were not generalizable to the Ontario setting. Therefore, we developed a primary economic evaluation for Ontario building on the methods of the INESSS analysis and other published economic studies.

Research Question

What is the cost-effectiveness of molecular testing compared to usual care for people with thyroid nodules of indeterminate cytology from the perspective of the Ontario Ministry of Health?

Methods

The information presented in this report follows the reporting standards set out by the Consolidated Health Economic Evaluation Reporting Standards (CHEERS) statement.[97]

Type of Analysis

We conducted a probabilistic cost–utility analysis because it is the reference case approach recommended by the Canadian Agency for Drugs and Technologies in Health (CADTH) guidelines for economic evaluation.[98] The effectiveness outcome is quality-adjusted life-years (QALYs), which considers both the person's survival and quality of life (e.g., 1 QALY represents 1 year of perfect health). A generic outcome measure such as the QALY allows decision-makers to make comparisons across different conditions and interventions.

We also conducted a cost-effectiveness analysis with outcomes expressed in natural units, including:

- Life-years
- Probability of predicting a correct diagnosis
- Rate of unnecessary surgeries (defined as benign nodules resected) in year 1

Target Population

Our target population was adults with a thyroid nodule of indeterminate cytology (TBSRTC III or IV). To be eligible for fine-needle aspiration biopsy, the nodule is usually 1 cm in diameter or greater on ultrasound. Based on information from the clinical evidence review of this health technology assessment, approximately 75% of the target population was female and the average age was approximately 55 years. We also conducted subgroup analyses for either TBSRTC III or IV alone.

Perspective

We conducted this analysis from the perspective of the Ontario Ministry of Health.

Intervention and Comparator

We compared the following strategies:

- Molecular testing (ThyroSeq v3 or Afirma GSC) to support the rule-out of suspected cancerous thyroid nodules prior to surgical removal (diagnostic lobectomy)
- Usual care:
  - Diagnostic lobectomy only (reference case analysis)
  - A proportion of patients underwent repeat fine-needle aspiration for further risk assessment, and the rest received diagnostic lobectomy (scenario analysis)

INTERVENTION

Although several tests are commercially available, we focused on the two tests that had been identified as having published analytical validation studies,[18] and on the most recent versions of those tests: ThyroSeq v3 and Afirma GSC. As outlined in the clinical evidence review section, we focused on the rule-out capacity of these tests.

When the result of the molecular test was conclusive (about 90% of the time[32,56]), we considered it to be dichotomous (positive or negative). For those with a positive result (both true and false positives), we obtained the surgical resection rate (diagnostic lobectomy) from the clinical evidence review of this health technology assessment (72% to 95%). We assumed that the remaining patients received surveillance. For those with a negative result (both true and false negatives), we also obtained the surgical resection rate (diagnostic lobectomy) from the clinical evidence review (6% to 7%). We assumed that the remaining patients received surveillance.

When the molecular test result was inconclusive (about 10% of the time, because of an insufficient sample[32,56]), we assumed that patients would receive a diagnostic lobectomy (based on the 2015 American Thyroid Association guidelines[9]).

COMPARATOR

For usual care, patients with a thyroid nodule of indeterminate cytology (TBSRTC III or IV) would be offered a repeat fine-needle aspiration or undergo diagnostic lobectomy. For nodules with a TBSRTC III cytology result, the current Cancer Care Ontario Diagnostic Pathway Map[99] recommends repeat fine-needle aspiration in 3 to 12 months or a second opinion on the cytology results. If the results of the repeat fine-needle aspiration are the same, patients can consider diagnostic lobectomy or surveillance (follow-up for at least 5 years). For nodules with a TBSRTC IV cytology result, diagnostic lobectomy is the long-established standard of care. However, there is regional variation in this standard, and many patients still opt for active surveillance.

For the reference case analysis, we considered diagnostic lobectomy to be the most appropriate comparator. This is because molecular tests are typically costly ($4,785 per test in Ontario), and if they were publicly funded, they would most likely be used by physicians to help decide if a patient needed diagnostic lobectomy. According to the recent INESSS recommendations[18] on molecular testing of thyroid nodules, “the use of molecular tests should be limited to patients whose test result is likely to influence management. Molecular testing should not be proposed to patients for whom surgery is indicated (high-risk factors), those for whom diagnostic lobectomy is not being considered because of

a low risk of malignancy, or those who express a preference for surgery or observation." In a scenario analysis, we considered usual care to be both repeat fine-needle aspiration and diagnostic lobectomy. We did not consider other advanced techniques used to improve preoperative diagnosis, because these are not routinely used in Ontario.

Time Horizon and Discounting

We used a 5-year time horizon for the reference case analysis to capture the effect of molecular testing on costs and outcomes (e.g., ongoing surveillance of thyroid nodules that were not surgically removed and management of permanent hypothyroidism and other complications related to thyroid surgeries). Because of a lack of long-term data on people with molecularly benign findings who did not undergo surgery (the available studies had median follow-up periods of less than 1 year), we did not use a lifetime horizon. We explored different time horizons in scenario analyses (10 and 20 years). In accordance with the CADTH guidelines,[98] we applied an annual discount rate of 1.5% to both costs and QALYs incurred after the first year.

Main Assumptions

The model's main assumptions were as follows:

- Although it is possible for an individual to have multiple thyroid nodules, for simplicity we assumed only one nodule per person. The average number of nodules per person was 1.05 in the primary studies, supporting this assumption (based on information from the clinical evidence review)
- For thyroid nodules of indeterminate cytology, the main goal of molecular testing is to correctly identify benign nodules, reduce the number of unnecessary diagnostic surgeries, and improve patients' quality of life (by avoiding surgery and surgery-related complications), while at the same time not missing any malignant nodules. The difference in cancer recurrence, progression, and survival (life-years) between patients who undergo molecular testing and patients who do not was likely to be very small. For simplicity, we did not model the possibility of cancer recurrence postsurgery or cancer progression in undetected malignant nodules. We also assumed that most benign nodules would remain stable with benign features during the model time horizon
- All patients included in the model were surgical candidates, had consented to surgery, and had no previous history of neck surgery or other confounding medical conditions
- To simplify the model, patients with histologically malignant nodules were assumed to be cured after diagnostic lobectomy and had the same mortality rate as the general population

Model Structure

We developed a model structure based on treatment pathways in Ontario (Cancer Care Ontario Diagnostic and Treatment Pathway Maps[99]), clinical practice guidelines (the 2015 American Thyroid Association guideline[9]), and published economic studies.[18,84,95] The model structure consisted of two parts: a decision tree (Figure 4) and Markov models (Figure 5).

For the molecular testing arm, we used a decision tree to calculate the proportions of patients with true positive, false positive, false negative, and true negative results according to test sensitivity, test specificity, and the prevalence of malignancy:

- If the result was positive (both true and false positives), about 83.5% of patients would receive a diagnostic lobectomy, and the rest would receive surveillance (based on information from the clinical evidence review)
- If the molecular test result was negative (both true and false negatives), about 6.5% of patients would receive a diagnostic lobectomy, and the rest would receive surveillance (based on information from the clinical evidence review)
- If the molecular test result was inconclusive (i.e., test failure), patients would receive a diagnostic lobectomy

For the usual care arm, all patients received diagnostic lobectomy.

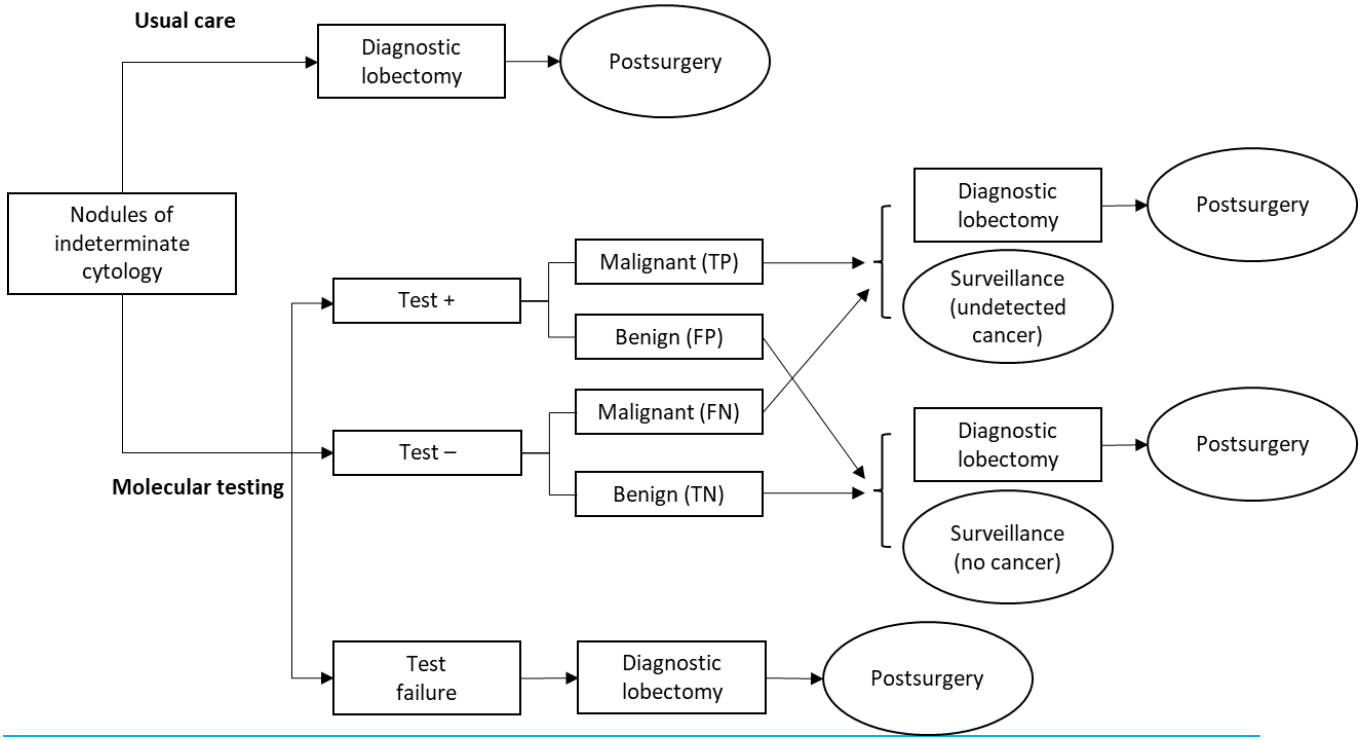

Figure 4: Model Structure—Decision Tree

Abbreviations: TP, true positive; FP, false positive; FN, false negative; TN, true negative.
Note: Boxes represent procedures or test results. Ovals represent mutually exclusive health states: postsurgery (Figure 5a); surveillance (no cancer; Figure 5b); surveillance (undetected cancer; Figure 5c).

Based on the decision tree, patients would then enter different Markov models depending on the treatment they received (surgery or surveillance) and whether they had undetected cancer. The cycle length of the Markov chains was 1 year, because patients are usually followed up once a year:

- Those who underwent diagnostic lobectomy would enter the postsurgery health state (Figure 5a). If the final histopathology result was benign, no further follow-up would be required; if the final histopathology result was malignant, the patient would be followed up annually by an endocrinologist for 5 years and then discharged to usual care with their family

physician (Antoine Eskander, email communication, March 28, 2021; Michael Odell, email communication, April 5, 2021). Approximately 22% of patients in this health state would have permanent hypothyroidism and require hormone replacement therapy.[13] During each annual cycle, patients could die from natural causes

- Those who underwent surveillance with no underlying cancer would enter the surveillance (no cancer) health state (Figure 5b). During each annual cycle, patients could die from natural causes
- Those who underwent surveillance with undetected cancer would enter the surveillance (undetected cancer) health state (Figure 5c). During each annual cycle, patients could remain well with a stable nodule, have thyroid cancer detected and undergo lobectomy, or die. Patients with undetected thyroid cancer would have a higher risk of death than those who were cancer-free. If a patient's cancer was detected later and they underwent surgery, they would enter the postsurgery health state

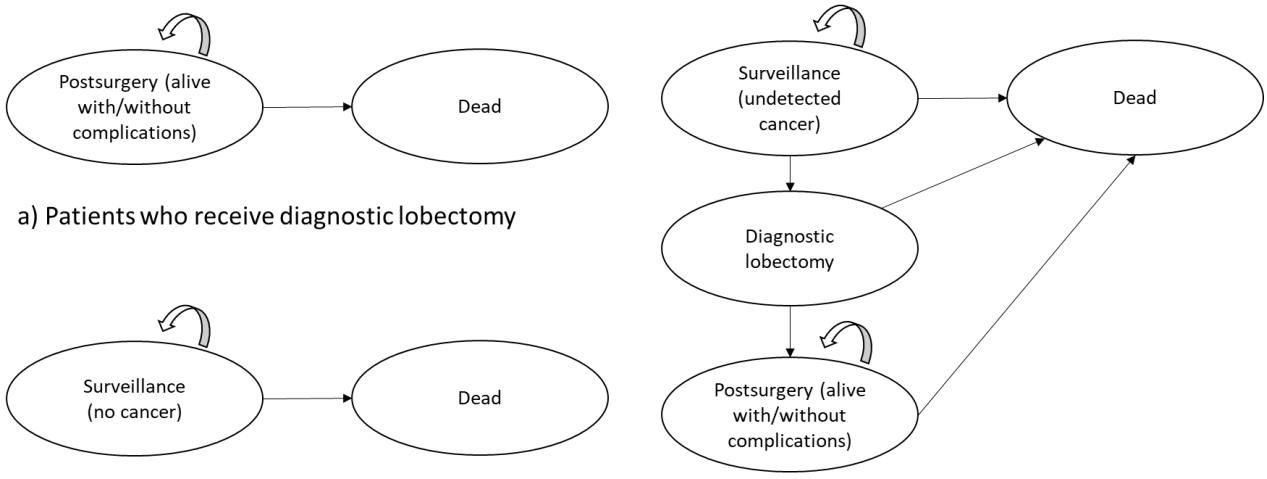

Figure 5: Model Structure—Markov Models

Abbreviations: TP, true positive; FP, false positive; FN, false negative; TN, true negative.
Note: Patients with undetected cancer (c) would have a higher risk of death than those who were cancer-free (a and b).

Clinical Parameters

The clinical parameters used in the model are shown in Table 7.

PREVALENCE OF MALIGNANCY

According to the literature, the prevalence (pretest probability) of malignancy in people with nodules of indeterminate cytology is approximately 25%.[17] The prevalence of malignancy in the Ontario target population likely varies from institution to institution; we used a wide range of possible values in our sensitivity analyses.

TEST PERFORMANCE

We obtained the sensitivity and specificity of the molecular tests (e.g., ThyroSeq v3, Afirma GSC) from two clinical validation studies[54,56] identified in the clinical evidence review of this health technology assessment.

PROBABILITY OF SURGICAL RESECTION

We obtained the probability of surgical resection after molecular testing from the clinical evidence review of this health technology assessment. The probability varied from study to study, but in general most patients with molecularly benign results would receive surveillance, and most patients with molecularly malignant results would undergo surgery.

PROBABILITY OF SURGICAL COMPLICATIONS OR SEQUELAE

We considered the cost and quality-of-life effects of key complications and sequelae associated with thyroid surgery. We included those that are frequent, severe, expensive to treat, or would have a great impact on the patient's health-related quality of life (e.g., hypothyroidism, hypoparathyroidism, recurrent laryngeal nerve injury). We also considered readmission and emergency department visits after surgery, and surgery-related death. We excluded complications that would have a negligible impact on health effects or resources.

PROBABILITY OF DEATH

For those who underwent diagnostic lobectomy and whose thyroid nodule was found to be malignant, we considered the surgery to be curative and assumed that patients would have the same mortality rate as the general population. We obtained the annual probability of all-cause natural mortality from the Canadian life table.[100] Patients with an undetected malignant thyroid nodule would be at higher risk of death (the relative risk of mortality for an undetected thyroid cancer is 1.8).

Table 7: Clinical Parameters for the Economic Model

Model parameter	Value (mean, 95% CI)	Distribution	Source
Patient characteristics			
Female, %	75%	Fixed	Clinical evidence review
Age, y	55	Fixed	Clinical evidence review
Prevalence of malignancy and test performance			
Prevalence (pretest probability) of malignancy	0.25 (0.1–0.40)	Beta	Cibas and Ali, 2017[17]
ThyroSeq v3 sensitivity	0.94 (0.86–0.98)	Beta	Clinical evidence review (Table 3)
ThyroSeq v3 specificity	0.82 (0.75–0.87)	Beta	Clinical evidence review (Table 3)
Afirma GSC sensitivity	0.91 (0.79–0.98)	Beta	Clinical evidence review (Table 3)
Afirma GSC specificity	0.68 (0.60–0.76)	Beta	Clinical evidence review (Table 3)
Probability of test failure	0.10 (0.05–0.15)	Beta	Steward et al, 2019[56]; Kay-Rivest et al, 2017[32]
Probability			
Surgical resection in case of molecularly suspect result (TP and FP)	83.5% (72%–95%)	Beta	Clinical evidence review (calculated based on data from Table 5)
Surgical resection in case of molecularly benign result (TN and FN)	6.5% (6%–7%)	Beta	Clinical evidence review (calculated based on data from Table 5)
Detecting thyroid cancer if nodules were not resected initially (annual; TP and FN)	20% (10%–33%)	Beta	Balentine et al, 2018[83]
Probability of surgical complications			
Permanent hypothyroidism after diagnostic lobectomy	22% (19%–27%)	Beta	Verloop et al, 2012[13]
Permanent hypothyroidism after total thyroidectomy	100%	Fixed	Expert opinion
Permanent hypoparathyroidism after total thyroidectomy	1%	Fixed	Al-Qurayshi et al, 2020[101]
Permanent recurrent laryngeal nerve injury after diagnostic lobectomy or total thyroidectomy	1%	Beta	Al-Qurayshi et al, 2020[101]
Readmission after diagnostic lobectomy	1.3%	Beta	Noel et al, 2021[102]

Model parameter	Value (mean, 95% CI)	Distribution	Source
Emergency department visit after diagnostic lobectomy	9.2%	Beta	Noel et al, 2021[102]
Surgery-related death	0.065%	Beta	Gómez-Ramírez et al, 2015[103]
Mortality			
Relative risk of mortality for patients with undetected thyroid cancer vs. the general population	1.80 (1.5–2.5)	Log-normal	INESSS, 2021[18]

Abbreviations: CI, confidence interval; FN, false negative; FP, false positive; GSC, gene sequencing classifier; INESSS, Institut national d'excellence en santé et services sociaux; TP, true positive; TN, true negative.

Utility Parameters

We considered the patient's health-related quality of life in the model. Utilities are numeric weights that represent a person's preference for a certain health state, such as postsurgery with no complications. Utilities are often measured on a scale from 0 (death) to 1 (perfect health). Disutilities represent the decrement in utility as a result of a particular symptom and are often expressed as negative values (e.g., −0.013 for a person undergoing diagnostic lobectomy). We obtained the utility and disutility parameters for the model from published economic studies (Table 8).[83,94] These utility values were estimated using a time trade-off methodology with a panel of thyroid experts.[94] We performed sensitivity analyses using alternative utility values.

Table 8: Utility Parameters for the Economic Model

Health state or treatment state	Utility/disutility (mean, 95% CI)	Adjusted utility[a]	Distribution	Source
Surveillance (no surgery)	0.98 (0.90 to 0.99)	0.909	Beta	Li et al, 2011[94]
Postsurgery, permanent hypothyroidism	0.94 (0.90 to 0.99)	0.869	Beta	Li et al, 2011[94]
Postsurgery, no complications	0.99 (0.90 to 0.99)	0.919	Beta	Li et al, 2011[94]
Postsurgery, permanent complications (e.g., recurrent laryngeal nerve injury or hypoparathyroidism)	0.70 (0.62 to 0.99)	0.629	Beta	Li et al, 2011[94]
Disutility due to lobectomy (1 week)	−0.013 (−0.057 to −0.005)	NA	Beta	Balentine et al, 2018[83]

Abbreviations: CI, confidence interval; NA, not applicable.

[a] Because the health state utility values were higher than those for the Canadian population in excellent health (a median of 0.919 for those aged 55 to 59 years old),[104] we adjusted these utilities downward by the difference.

Cost Parameters

Cost parameters we used in the model are presented in Table 9. We considered the following types of costs in our model:

- Cost of an additional fine-needle aspiration biopsy, a physician visit, and counselling associated with molecular testing
- Hospital cost of surgery (e.g., diagnostic lobectomy as an inpatient procedure or day surgery)
- Professional fees for surgery (e.g., surgeon, surgical assistant, and anesthesiologist)
- Cost of molecular testing (e.g., sample collection, shipping, processing, testing, interpretation, and reporting)
- Cost of long-term surveillance
    - Postsurgery
    - People who received molecular testing but did not undergo surgery (majority with nodules that were molecularly benign, some molecularly malignant)
- Cost of treating surgical complications and sequelae (short- and long-term)

We obtained cost inputs from standard Ontario sources, published literature, and clinical experts. We obtained fees for professional visits and procedures from the Ontario Schedule of Benefits for Physician Services.[105] We obtained hospital costs from the Ontario Case Costing database[96] of the Ministry of Health. We obtained diagnostic and laboratory fees from the Ontario Schedule of Benefits for Laboratory Services.[106] We obtained costs related to molecular testing from the manufacturers. Because the average age of our target population was less than 65 years, drug costs would not usually be covered by the Ontario Health Insurance Plan, so we excluded drug costs from the model. All costs are reported in 2021 Canadian dollars. When values in 2021 Canadian dollars were not available, we used the health care component of the Statistics Canada Consumer Price Index[107] to adjust the costs to 2021 Canadian dollars.

TREATMENT OF SURGICAL COMPLICATIONS AND SEQUELAE

Hypothyroidism occurs in about 22% of those who undergo diagnostic lobectomy and 100% of those who undergo total thyroidectomy. For management of hypothyroidism, we assumed that people would see an endocrinologist once a year and receive thyroid hormone replacement therapy. Patients would also have thyroid function tests 2 to 3 months post-treatment to check the adequacy of the thyroid hormone replacement therapy.

Permanent hypoparathyroidism occurs in about 1% of people who undergo total thyroidectomy (it is not possible for people who undergo diagnostic lobectomy). For management of hypoparathyroidism, we assumed that people would see an endocrinologist once a year and receive oral calcium carbonate and calcitriol. They would also have blood tests twice a year to monitor calcium and phosphorus levels.

Permanent injuries to the recurrent laryngeal nerve occur in about 1% of people who undergo diagnostic lobectomy or total thyroidectomy. For management of recurrent laryngeal nerve injury, we assumed that people would see a laryngologist and receive an injection laryngoplasty.

COST OF LONG-TERM SURVEILLANCE

For patients who did not undergo surgery or patients whose nodules were found to be malignant after diagnostic lobectomy, we assumed that they would receive surveillance from an endocrinologist for 5 years (annual physical examination, blood work, neck ultrasound) and then be discharged to their primary care provider.

Table 9: Cost Parameters for the Economic Model

Model parameter	Mean cost, $[a]	Distribution	Source
Molecular test	4,785.00	Fixed	GROUP Thyroid Molecular Testing and Coordination, email communication, May 11, 2021
Additional physician visit associated with molecular testing	44.40	Fixed	OHIP Schedule of Benefits A033[105]
Ultrasound-guided fine-needle aspiration biopsy (for molecular testing or repeat fine-needle aspiration)	162.13	—	Calculated
Physician fees	150.10	Fixed	OHIP Schedule of Benefits J243, Z771, J105B, and J105C[105]
Laboratory fees	12.03	Fixed	OHIP Schedule of Benefits L705[106]
Diagnostic lobectomy	5,412.06	—	Calculated
Preprocedural assessment and laboratory tests	49.14	Fixed	OHIP Schedule of Benefits A033, L341, and L045[105,106]
Periprocedural hospital cost	4,346.65	Gamma	Ontario Case Costing Initiative, 2018[96,b]
Periprocedural physician fees[c]	971.87	Fixed	OHIP Schedule of Benefits[105]; assumed operation takes 1.5 h[d]
Postprocedural assessment and laboratory tests	44.40	Fixed	OHIP Schedule of Benefits A033[105]
Total thyroidectomy	7,335.62	—	
Preprocedural assessment and laboratory tests	49.14	Fixed	OHIP Schedule of Benefits A033, L341, and L045[105,106]
Periprocedural hospital cost	5,754.42	Gamma	Ontario Case Costing Initiative, 2018[96,e]
Periprocedural physician fees[c]	1,364.76	Fixed	OHIP Schedule of Benefits[105]; assumed operation takes 2 h[d]
Postprocedural assessment and laboratory tests	167.30	Fixed	OHIP Schedule of Benefits A033, A153, Z296, L330, and L045[105,106]
Final surgical pathology on thyroid specimen	18.75	Fixed	OHIP Schedule of Benefits L720[106]
Readmission after thyroid surgery	5,663.47	Fixed	CIHI 2020[108]
Emergency department visit after thyroid surgery	309.40	Fixed	CIHI 2020[109]

Model parameter	Mean cost, $[a]	Distribution	Source
Annual surveillance (nodules that are not resected and nodules found to be malignant after diagnostic surgery)	169.11	Fixed	OHIP Schedule of Benefits A153 (medical-specific assessment with an endocrinologist), J105B and J105C (ultrasound), L341 (TSH), and L609 (serum thyroglobulin)[105,106]
Annual treatment for hypothyroidism	89.91	Fixed	OHIP Schedule of Benefits A153, L607, L339, and L341[105,106]
Annual treatment for hypoparathyroidism	87.39	Fixed	OHIP Schedule of Benefits A153, L045, and L194[105,106]
Treatment for recurrent laryngeal nerve injury	1,023.08	Fixed	Tam et al, 2017[110]

Abbreviations: CIHI, Canadian Institute for Health Information; OHIP, Ontario Health Insurance Plan; TSH, thyroid-stimulating hormone.

[a] In 2021 Canadian dollars.

[b] Calculated as the weighted average cost of an inpatient procedure (74%, $4,630 in 2018 CAD) and a day surgery (26%, $3,041 in 2018 CAD).

[c] Included professional fees for surgeon, surgical assistant, and anaesthesiologist.

[d] Estimated based on clinical expert opinion (Antoine Eskander, email communication, March 28, 2021; Michael Odell, email communication, April 5, 2021).

[e] Calculated as the weighted average cost of an inpatient procedure (94%, $5,728 in 2018 CAD) and a day surgery (6%, $3,459 in 2018 CAD).

Internal Validation

Formal internal validation was conducted by a secondary health economist. This included testing the mathematical logic of the model and checking for errors and accuracy of parameter inputs and equations.

Analysis

We conducted a reference case analysis and sensitivity analyses. Our reference case analysis adhered to the CADTH guidelines[98] when appropriate. The reference case represents the analysis with the most likely set of input parameters and model assumptions. Our sensitivity analyses explored how the results would be affected by varying input parameters and model assumptions.

For the reference case analysis, we conducted a probabilistic analysis to capture parameter uncertainty. When possible, we specified distributions around input parameters using the mean and standard error. Selected cost parameters were characterized by gamma distributions; probabilities and utilities were characterized by beta distributions; and relative risks were characterized by log-normal distributions. We ran a total of 5,000 simulations and calculated the expected values of costs and outcomes for each strategy. We presented the probability that each strategy was cost-effective over a range of willingness-to-pay values on a cost-effectiveness acceptability curve.

We also examined additional structural and parameter uncertainty by conducting several scenario analyses (Table 10). In one scenario analysis, we widened the criteria for people who were eligible for molecular testing and compared molecular testing to an alternative comparator: usual care consisted of repeat fine-needle aspiration biopsy and diagnostic lobectomy. In the reference case, we assumed that molecular testing would be used only to help physicians decide if a patient needed diagnostic surgery (therefore, the comparator was diagnostic lobectomy only). However, according to recent guidelines,[9] patients with nodules of indeterminate cytology can receive either repeat fine-needle aspiration biopsy or diagnostic lobectomy. For this scenario, we estimated the percentage of repeat fine-needle aspiration biopsies among patients with nodules of indeterminate cytology using the Ontario administrative data (approximately 23.3%). Based on a recent Canadian study,[22] approximately 40% of nodules of indeterminate cytology remained indeterminate after a repeat fine-needle aspiration biopsy; 51% were reclassified as nondiagnostic or benign; and 9% were reclassified as malignant. We assumed that patients with a nondiagnostic or benign cytology result would receive surveillance; patients with a malignant cytology result would receive total thyroidectomy; and the repeat fine-needle aspiration result could be considered accurate.

Table 10: Variables Varied in Scenario Analyses

Parameter or assumption	Reference case	Scenario analysis
Comparator	Usual care consisted of only diagnostic lobectomy	Usual care consisted of repeat fine-needle aspiration biopsy (23.3%) and diagnostic lobectomy
Prevalence of malignancy	Average value based on Cibas and Ali[17]: 0.25	Range of values based on Bernstein et al, 2016[23]: lower estimate: 0.06; upper estimate: 0.46
Molecular test sensitivity	Afirma GSC: 91% ThyroSeq v3: 94%	Upper and lower 95% CI
Molecular test specificity	Afirma GSC: 68% ThyroSeq v3: 82%	Upper and lower 95% CI
Probability of surgical resection in people with a molecularly benign result	Based on clinical studies of molecular tests: 6%–7%	0% or 20%
Probability of surgical resection in people with a molecularly malignant result	Based on clinical studies of molecular tests: 72%–95%	60% or 100%
Utility parameters	Utility parameters based on the INESSS analysis[18] and Li et al, 2011[94] (higher utility for postsurgery vs. surveillance health states: 0.99 vs. 0.98)	Assumptions: • Equal utility for postsurgery vs. surveillance health states: 0.99 vs. 0.99 • Lower utility for postsurgery vs. surveillance health states: 0.98 vs. 0.99
Cost of fine-needle aspiration biopsy	Assumed that one extra fine-needle aspiration biopsy was needed to obtain sample for molecular testing	Assumed that the initial fine-needle aspiration biopsy sample was used for molecular testing
Cost of molecular test	$4,785 per test	± 25%
Cost of surveillance	$169 per year	Two times higher
TBSRTC III nodules only (changing the sensitivity and specificity only)	Afirma: sensitivity 91%, specificity 68% ThyroSeq: sensitivity 94%, specificity 82%	Afirma: sensitivity 93%, specificity 71% ThyroSeq: sensitivity 91%, specificity 85%
TBSRTC IV nodules only (changing the sensitivity and specificity only)	Afirma: sensitivity 91%, specificity 68% ThyroSeq: sensitivity 94%, specificity 82%	Afirma: sensitivity 88%, specificity 64% ThyroSeq: sensitivity 97%, specificity 75%
Time horizon	5 years	10 years; 20 years
Discount rate	1.5%	0%, 3%

Abbreviations: CI, confidence interval; GSC, gene sequencing classifier; INESSS, Institut national d'excellence en santé et services sociaux; QALY, quality-adjusted life-years; TBSRTC, The Bethesda System for Reporting Thyroid Cytopathology.

Results

Reference Case Analysis

Tables 11 and 12 present the results of the reference case analysis. The average total cost per patient was $8,503 for Afirma GSC, $8,152 for ThyroSeq v3, and $5,838 for diagnostic lobectomy. Compared to diagnostic lobectomy, molecular testing led to increased costs associated with the test itself ($4,785), an additional fine-needle aspiration biopsy and physician visit ($207), and additional surveillance and follow-up for nodules that were not resected ($370 to $431). Molecular testing led to reduced costs related to surgery and pathology ($2,594 to $2,991) and treatment for surgical complications and sequelae ($103 to $118), but these cost reductions were not enough to offset the cost increases. Overall, adding molecular testing to the diagnostic pathway for nodules of indeterminate cytology would increase the total costs by $2,313 to $2,664 per patient.

Compared to diagnostic lobectomy, molecular testing led to minimal change in life-years (0.0002 to 0.0005 over a 5-year time horizon) and a slight improvement in QALYs (0.0089 to 0.0105). However, it increased the probability of predicting a correct diagnosis (from 25.2% to 67.3%–77.2%), reduced the probability of surgery (from 100% to 42.3%–49.3%), and reduced the probability of unnecessary surgery (from 74.8% to 21.2%–28.6%). The resulting ICERs were $220,572 to $298,653 per QALY gained, $4,451 to $6,328 per additional correct diagnosis, and $4,314 to $5,769 per unnecessary surgery avoided.

Table 11: Reference Case Analysis Results—Afirma GSC Versus Diagnostic Lobectomy

Finding	Afirma GSC, mean (95% CrI)	Diagnostic lobectomy, mean (95% CrI)	Difference
Average total cost	$8,502.87 ($7,898.14 to $9,197.85)	$5,838.46 ($5,312.48 to $6,394.09)	$2,664.42 ($2,101.02 to $3,206.71)
Cost of molecular test	$4,785.00	$0.00	$4,785.00
Cost of fine-needle aspiration and additional physician visit	$206.53	$0.00	$206.53
Cost of surgery and pathology	$2,832.38	$5,426.72	−$2,594.34
Cost of surveillance or follow-up	$574.87	$204.67	$370.20
Cost of surgical complications and sequelae	$104.09	$207.07	−$102.98
Average life-years	4.8011 (4.7990 to 4.8024)	4.8009 (4.7992 to 4.8022)	0.0002 (−0.0018 to 0.0017)
Average QALYs	4.3802 (4.2364 to 4.5027)	4.3712 (4.1795 to 4.5220)	0.0089 (−0.1261 to 0.1388)
Probability of making a correct diagnosis	67.3% (59.5% to 74.7%)	25.2% (11.8% to 41.4%)	42.1% (27.8% to 54.8%)
Benign nodules (true negative)	46.1%	0.0%	46.1%
Malignant nodules (true positive)	21.2%	25.2%	−4.0%
Probability of unnecessary surgeries in year 1	28.6% (21.0% to 36.8%)	74.8% (58.6% to 88.2%)	−46.2% (−56.7% to −35.1%)
ICER (cost per QALY gained)	–	–	$298,653
ICER (cost per additional correct diagnosis)	–	–	$6,328
ICER (cost per unnecessary surgery avoided)	–	–	$5,769

Abbreviations: CrI, credible interval; ICER, incremental cost-effectiveness ratio; QALY, quality-adjusted life-years.

Table 12: Reference Case Analysis Results—ThyroSeq v3 Versus Diagnostic Lobectomy

Finding	ThyroSeq v3, mean (95% CrI)	Diagnostic lobectomy, mean (95% CrI)	Difference
Average total cost	$8,151.95 ($7,535.14 to $8,867.42)	$5,838.46 ($5,312.48 to $6,394.09)	$2,313.49 ($1,724.73 to $2,918.43)
Cost of molecular test	$4,785.00	$0.00	$4,785.00
Cost of fine-needle aspiration and additional physician visit	$206.53	$0.00	$206.53
Cost of surgery and pathology	$2,435.46	$5,426.72	−$2,991.26
Cost of surveillance or follow-up	$635.70	$204.67	$431.03
Cost of surgical complications and sequelae	$89.26	$207.07	−$117.81
Average life-years	4.8014 (4.7994 to 4.8026)	4.8009 (4.7992 to 4.8022)	0.0005 (−0.0014 to 0.0020)
Average QALYs	4.3817 (4.2328 to 4.5085)	4.3712 (4.1795 to 4.5220)	0.0105 (−0.1447 to 0.1621)
Probability of making a correct diagnosis	77.2% (70.8% to 83.2%)	25.2% (11.8% to 41.4%)	52.0% (36.7% to 65.2%)
Benign nodules (true negative)	55.5%	0.0%	55.5%
Malignant nodules (true positive)	21.7%	25.2%	−3.5%
Probability of unnecessary surgeries in year 1	21.2% (15.2% to 27.6%)	74.8% (58.6% to 88.2%)	−53.6% (−64.4% to −41.3%)
ICER (cost per QALY gained)	–	–	$220,572
ICER (cost per additional correct diagnosis)	–	–	$4,451
ICER (cost per unnecessary surgery avoided)	–	–	$4,314

Abbreviations: CrI, credible interval; ICER, incremental cost-effectiveness ratio; QALY, quality-adjusted life-years.

Figures 6 and 7 present cost-effectiveness acceptability curves, which show the probability of each strategy being cost-effective across a range of willingness-to-pay values (see Appendix 6 for the cost-effectiveness plane scatterplot). At the commonly used willingness-to-pay values of $50,000 and $100,000 per QALY, molecular testing was moderately likely not to be cost-effective (i.e., the probability of molecular testing being cost-effective was less than 50%).

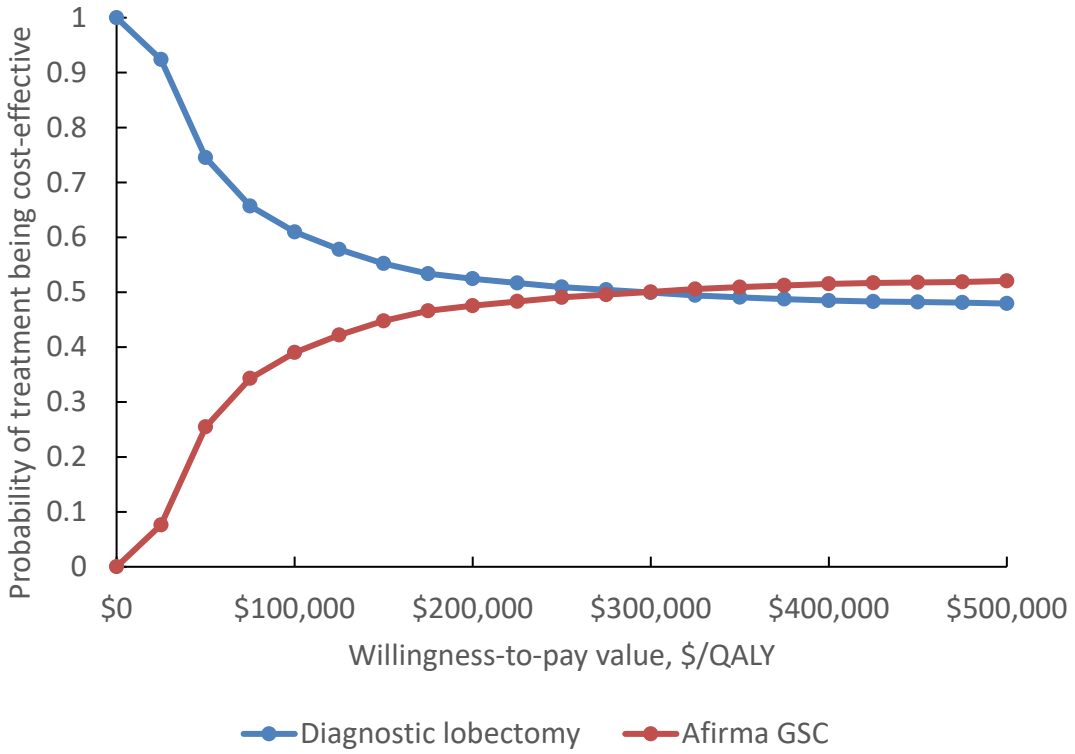

Figure 6: Cost-Effectiveness Acceptability Curve—Afirma GSC Versus Diagnostic Lobectomy

Abbreviation: GSC, gene sequencing classifier; QALY, quality-adjusted life-years.

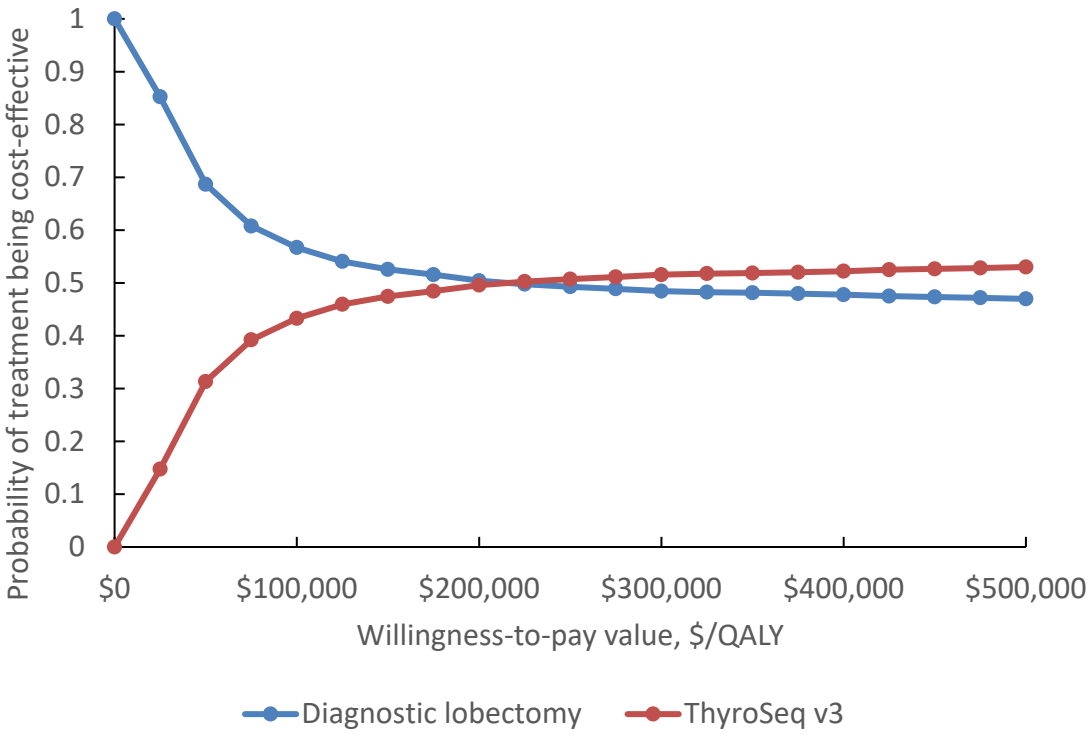

Figure 7: Cost-Effectiveness Acceptability Curve–ThyroSeq v3 Versus Diagnostic Lobectomy

Abbreviation: QALY, quality-adjusted life-years.

Scenario Analyses

The results of the scenario analyses are shown in Tables 13 and 14. Molecular testing was more costly and more effective in all scenarios. The ICERs ranged from $37,505 per QALY to $539,325 per QALY. The ICER decreased substantially when we assumed a lower prevalence of malignancy in the target population (scenario 2a); a higher test specificity (scenario 4a); an equal or lower utility for postsurgery versus surveillance health states (scenarios 7a and 7b); a lower cost for the molecular test (scenario 9b); a higher cost for surgery (scenario 11); and a longer time horizon (scenarios 13a and 13b). The ICER increased substantially when we assumed a higher prevalence of malignancy in the target population (scenario 2b); a lower test specificity (scenario 4b); and a higher cost for the molecular test (scenario 9a). We also conducted subgroup analyses for thyroid nodules classified as TBSRTC III only or TBSRTC IV only, by varying the test sensitivity and specificity. The ICER was lower in the TBSRTC III–only subgroup (scenario 12a) but higher in the TBSRTC IV–only subgroup (scenario 12b).

We conducted two scenario analyses using a longer time horizon (scenarios 13a and 13b). In these scenarios, we assumed that people with molecularly benign results were followed for only 5 years, not the entire model time horizon. Molecular testing became more cost-effective in these scenarios because the long-term quality-of-life benefit accumulated but the long-term surveillance cost did not. However, if we assumed that people with molecularly benign results were followed for a longer period, molecular testing would be less cost-effective because the long-term surveillance costs would accumulate.

Table 13: Scenario Analysis Results—Afirma GSC Versus Diagnostic Lobectomy

Scenario	Total cost			Total QALYs			ICER ($/QALY)	% Change vs. reference case
	Afirma GSC	Diagnostic lobectomy	Difference	Afirma GSC	Diagnostic lobectomy	Difference		
Reference case	$8,502.87	$5,838.46	$2,664.42	4.3802	4.3712	0.0089	$298,653.37	–
1: Usual care consisted of repeat fine-needle aspiration and diagnostic lobectomy	$8,514.63	$5,359.53	$3,155.09	4.3810	4.3709	0.0101	$311,713.77	+4%
2a: Prevalence of malignancy, 6%	$7,855.94	$5,677.50	$2,178.44	4.3808	4.3704	0.0104	$210,204.42	−30%
2b: Prevalence of malignancy, 46%	$9,205.92	$6,001.09	$3,204.83	4.3787	4.3726	0.0061	$523,203.05	+75%
3a: Sensitivity, upper 95% CI	$8,521.98	$5,824.79	$2,697.19	4.3838	4.3739	0.0098	$273,965.46	−8%
3b: Sensitivity, lower 95% CI	$8,462.47	$5,842.63	$2,619.85	4.3826	4.3743	0.0082	$318,852.22	+7%
4a: Specificity, upper 95% CI	$8,316.38	$5,839.53	$2,476.85	4.3833	4.3715	0.0118	$209,648.15	−30%
4b: Specificity, lower 95% CI	$8,744.92	$5,849.70	$2,895.22	4.3778	4.3704	0.0074	$391,246.96	+31%
5a: Probability of surgical resection in people with molecularly benign result, 0%	$8,358.70	$5,833.39	$2,525.32	4.3832	4.3738	0.0094	$267,309.44	−10%
5b: Probability of surgical resection in people with molecularly benign result, 20%	$8,806.35	$5,843.86	$2,962.49	4.3776	4.3683	0.0093	$320,130.95	+7%
6a: Probability of surgical resection in people with molecularly malignant result, 60%	$8,122.95	$5,833.80	$2,289.15	4.3833	4.3747	0.0087	$264,261.84	−12%
6b: Probability of surgical resection in people with molecularly malignant result, 100%	$8,727.50	$5,843.04	$2,884.46	4.3805	4.3721	0.0084	$341,716.27	+14%
7a: Utility parameters: equal utility for postsurgery and surveillance	$8,510.52	$5,837.85	$2,672.67	4.4068	4.3750	0.0319	$83,849.37	−72%
7b: Utility parameters: lower utility for postsurgery vs. surveillance	$8,512.69	$5,844.06	$2,668.64	4.3849	4.3317	0.0532	$50,180.38	−83%
8: Cost of an extra fine-needle aspiration biopsy not needed	$8,347.40	$5,838.70	$2,508.71	4.3810	4.3720	0.0090	$280,041.44	−6%
9a: Cost of molecular test, +25%	$9,692.65	$5,836.12	$3,856.53	4.3786	4.3714	0.0072	$539,325.77	+81%
9b: Cost of molecular test, −25%	$7,307.44	$5,838.12	$1,469.32	4.3787	4.3696	0.0091	$161,334.85	−46%
10: Cost of surveillance, 2× higher	$9,068.30	$6,043.53	$3,024.77	4.3811	4.3709	0.0102	$296,451.11	−1%
11: Cost of surgery, +25%	$9,229.49	$7,209.95	$2,019.55	4.3807	4.3719	0.0087	$231,008.85	−23%

Scenario	Total cost			Total QALYs			ICER ($/QALY)	% Change vs. reference case
	Afirma GSC	Diagnostic lobectomy	Difference	Afirma GSC	Diagnostic lobectomy	Difference		
12a: TBSRTC III nodules only	$8,440.33	$5,848.76	$2,591.57	4.3853	4.3755	0.0098	$264,421.19	−11%
12b: TBSRTC IV nodules only	$8,596.03	$5,840.27	$2,755.76	4.3807	4.3732	0.0074	$370,725.06	+24%
13a: Time horizon, 10 years	$8,621.49	$5,932.29	$2,689.20	8.3478	8.3283	0.0195	$138,245.86	−54%
13b: Time horizon, 20 years	$8,709.81	$6,059.91	$2,649.90	14.9764	14.9567	0.0197	$134,753.68	−55%
14a: Discount rate, 0%	$8,517.44	$5,850.41	$2,667.03	4.5118	4.5017	0.0101	$265,329.38	−11%
14b: Discount rate, 3%	$8,496.44	$5,837.66	$2,658.78	4.2589	4.2495	0.0094	$282,066.49	−6%

Abbreviations: CI, confidence interval; GSC, gene sequencing classifier; ICER, incremental cost-effectiveness ratio; QALY, quality-adjusted life-year; TBSRTC, The Bethesda System for Reporting Thyroid Cytopathology.

Table 14: Scenario Analysis Results—ThyroSeq v3 Versus Diagnostic Lobectomy

Scenario	Total cost			Total QALYs			ICER ($/QALY)	% Change vs. reference case
	ThyroSeq v3	Diagnostic lobectomy	Difference	ThyroSeq v3	Diagnostic lobectomy	Difference		
Reference case	$8,151.95	$5,838.46	$2,313.49	4.3817	4.3712	0.0105	$220,571.50	–
1: Usual care consisted of repeat fine-needle aspiration and diagnostic lobectomy	$8,159.88	$5,359.53	$2,800.34	4.3824	4.3709	0.0116	$242,095.72	+10%
2a: Prevalence of malignancy, 6%	$7,403.92	$5,677.50	$1,726.43	4.3827	4.3704	0.0122	$141,233.11	−36%
2b: Prevalence of malignancy, 46%	$8,962.65	$6,001.09	$2,961.56	4.3800	4.3726	0.0075	$395,711.63	+79%
3a: Sensitivity, upper 95% CI	$8,157.68	$5,824.79	$2,332.89	4.3852	4.3739	0.0113	$206,457.32	−6%
3b: Sensitivity, lower 95% CI	$8,124.47	$5,842.63	$2,281.85	4.3845	4.3743	0.0101	$225,533.69	+2%
4a: Specificity, upper 95% CI	$8,043.67	$5,839.53	$2,204.14	4.3849	4.3715	0.0133	$165,356.30	−25%
4b: Specificity, lower 95% CI	$8,373.09	$5,849.70	$2,523.39	4.3793	4.3704	0.0089	$282,063.68	+28%
5a: Probability of surgical resection in people with molecularly benign result, 0%	$7,976.95	$5,833.39	$2,143.56	4.3853	4.3738	0.0115	$186,289.29	−16%
5b: Probability of surgical resection in people with molecularly benign result, 20%	$8,511.95	$5,843.86	$2,668.09	4.3792	4.3683	0.0109	$244,687.06	+11%
6a: Probability of surgical resection in people with molecularly malignant result, 60%	$7,882.53	$5,833.80	$2,048.73	4.3844	4.3747	0.0097	$211,008.50	−4%
6b: Probability of surgical resection in people with molecularly malignant result, 100%	$8,302.75	$5,843.04	$2,459.72	4.3825	4.3721	0.0104	$237,306.81	+8%
7a: Utility parameters: equal utility for postsurgery and surveillance	$8,162.47	$5,837.85	$2,324.61	4.4118	4.3750	0.0368	$63,098.35	−71%
7b: Utility parameters: lower utility for postsurgery vs. surveillance	$8,152.98	$5,844.06	$2,308.92	4.3932	4.3317	0.0616	$37,505.21	−83%
8: Cost of an extra fine-needle aspiration biopsy not needed	$7,994.25	$5,838.70	$2,155.55	4.3826	4.3720	0.0105	$205,129.69	−7%
9a: Cost of molecular test, +25%	$9,344.48	$5,836.12	$3,508.36	4.3800	4.3714	0.0086	$409,397.35	+86%
9b: Cost of molecular test, −25%	$6,957.34	$5,838.12	$1,119.22	4.3804	4.3696	0.0107	$104,168.26	−53%
10: Cost of surveillance, 2× higher	$8,773.54	$6,043.53	$2,730.00	4.3827	4.3709	0.0118	$231,514.04	+5%

Scenario	Total cost			Total QALYs			ICER ($/QALY)	% Change vs. reference case
	ThyroSeq v3	Diagnostic lobectomy	Difference	ThyroSeq v3	Diagnostic lobectomy	Difference		
11: Cost of surgery, +25%	$8,774.46	$7,209.95	$1,564.51	4.3821	4.3719	0.0102	$153,445.96	−30%
12a: TBSRTC III nodules only	$8,059.82	$5,848.76	$2,211.06	4.3866	4.3755	0.0111	$199,287.45	−10%
12b: TBSRTC IV nodules only	$8,351.14	$5,840.27	$2,510.87	4.3820	4.3732	0.0088	$284,840.32	+29%
13a: Time horizon, 10 years	$8,262.05	$5,932.29	$2,329.76	8.3516	8.3283	0.0233	$99,994.76	−55%
13b: Time horizon, 20 years	$8,329.12	$6,059.91	$2,269.21	14.9826	14.9567	0.0259	$87,687.38	−60%
14a: Discount rate, 0%	$8,174.15	$5,850.41	$2,323.74	4.5136	4.5017	0.0118	$196,215.92	−11%
14b: Discount rate, 3%	$8,144.83	$5,837.66	$2,307.17	4.2604	4.2495	0.0110	$210,145.60	−5%

Abbreviations: CI, confidence interval; GSC, gene sequencing classifier; ICER, incremental cost-effectiveness ratio; QALY, quality-adjusted life-year; TBSRTC, The Bethesda System for Reporting Thyroid Cytopathology.

Threshold Analyses

We also conducted threshold analyses to determine at what list price molecular testing would become less costly than diagnostic lobectomy (i.e., the break-even point). We found that Afirma GSC would need to cost less than $2,150 per test, and ThyroSeq v3 would need to cost less than $2,488 per test.

Discussion

Our results showed that molecular testing could improve the probability of predicting a correct diagnosis, reduce unnecessary surgeries, and slightly improve patients' QALYs, but it would increase the average total cost by about $2,300 to $2,700 per patient. Incorporating molecular testing into the diagnostic pathway would add an initial cost of $4,785 per test, but it could lead to downstream cost savings if unnecessary surgeries could be avoided in patients who did not have a malignancy. However, molecular testing could also increase long-term costs because patients who do not undergo surgery would require ongoing surveillance, and these costs could accumulate over time. Because there was only a slight improvement in QALYs but a large difference in costs for molecular testing, the ICER compared to diagnostic lobectomy was $220,572 to $298,653 per QALY gained. Probabilistic analyses suggested that at commonly used willingness-to-pay values of $50,000 and $100,000 per QALY, molecular testing was unlikely to be cost-effective compared with diagnostic lobectomy.

Our results were consistent with those of published economic studies in terms of the effectiveness outcomes. We had almost identical results to the study by Nicholson et al,[95] which estimated the probability of predicting a correct diagnosis (25.0% for diagnostic lobectomy, 73.2% for ThyroSeq v3, and 63.7% for Afirma GSC), because key clinical input parameters (sensitivity, specificity, and the prevalence of malignancy) were the same (in our analysis, the probability of predicting a correct diagnosis was 25.2% for diagnostic lobectomy, 77.2% for ThyroSeq v3, and 67.3% for Afirma GSC). Compared to the INESSS study,[18] which estimated QALYs over a 5-year time horizon, we also found that molecular testing could lead to a slight improvement in QALYs compared to diagnostic lobectomy (about 0.01 over a 5-year time horizon). Several previous cost–utility studies of molecular testing (older versions of the tests)[83,87,92,93] also found that the difference in QALYs between molecular testing and standard care was minimal. There was no difference in life expectancy across strategies, likely because thyroid surgery is a relatively safe procedure (low risks of surgery-related death and complications), and because differentiated thyroid cancers are generally indolent and missed cancers rarely cause substantial morbidity and mortality. The small difference in QALYs between strategies was driven mainly by differences in utilities between the postsurgery and surveillance health states, because the main effect of molecular testing was to reduce the number of surgeries.

Our results were different from previous economic studies in terms of costs because resource use and cost parameters were different for each setting. Similar to the study by Nicholson et al,[95] we found that molecular testing would increase the average total cost per patient. However, the INESSS study[18] found ThyroSeq v3 to be $37 less costly and Afirma GSC to be $272 more costly compared to diagnostic lobectomy. This was likely because of a larger cost difference between molecular testing and diagnostic lobectomy in Quebec ($5,385 for molecular testing vs. $9,842 for diagnostic lobectomy) compared to Ontario ($4,784 for molecular testing vs. $5,412 for diagnostic lobectomy).

Our scenario analyses suggested that the cost-effectiveness of molecular testing compared to diagnostic lobectomy was strongly affected by the utility parameter values. As pointed out in several published studies,[94] there are some limitations with health state utility parameters. Robust utility estimates (that

have been directly elicited from patients with nodules of indeterminate cytology who experienced surveillance, thyroid surgery, and its complications) are lacking. Instead, most utility values used by published economic studies were estimated by clinical experts, who usually suggested the postsurgery health state to be preferable to the surveillance (no surgery) health state. As a result, patients in the molecular testing arm would have had fewer QALYs, because more patients would be in the surveillance (no surgery) health state if we did not consider the QALY loss caused by surgical complications and sequelae. Some studies have also suggested that physicians and patients may have discordant impressions of patients' quality of life in various health states.[111] It is possible that some patients may experience increased anxiety with annual surveillance, and thyroid surgery could provide "closure" and eliminate the need for ongoing follow-up visits. However, it is also possible that some patients would experience decreased quality of life after surgery and would prefer surveillance. To better understand the cost-effectiveness of molecular testing, further research into the utilities of different health states in this patient population is needed. Our scenario analyses showed that when we assumed that the surveillance (no surgery) health state was seen as preferable or equal to the postsurgery health state, molecular testing became more cost-effective.

The cost-effectiveness of molecular testing was also very sensitive to variations in the prevalence of malignancy in patients with nodules of indeterminate cytology. This parameter was related to how well the test could perform as a "rule-out" test. The rate of malignancy usually varies from one institution to another based on cytologic expertise and other factors (e.g., high-volume centres may tend to have higher rates of malignancy).[112] When we assumed a lower rate of malignancy (6%) while keeping the test sensitivity and specificity constant, the ICER decreased, suggesting that molecular testing would be more cost-effective as a "rule-out" test in such a setting. When we assumed a higher rate of malignancy (46%), the ICER increased significantly, suggesting that molecular testing would not be cost-effective in such a setting.

Finally, although we found that molecular testing has the potential to decrease the probability of surgery, some published studies (in older versions of the test) have suggested that the availability of molecular testing may increase the number of indeterminate fine-needle aspiration diagnoses (i.e., thyroid nodules are more likely to be classified as being of indeterminate cytology). As a result, the institutional rates of surgery and malignancy may not change, raising uncertainty about the benefits of risk stratification with molecular testing.[113]

Strengths and Limitations

Our analysis had several strengths. First, it evaluated the most recent versions of the molecular tests and used Ontario-specific inputs for resource use and costs. Second, although the QALY is usually the recommended outcome measure for economic evaluations, we assessed the effectiveness of molecular testing using several different outcomes specific to the health condition, including life-years, the probability of predicting a correct diagnosis, the probability of surgery, and the probability of unnecessary surgery. These additional outcomes may help provide a clearer picture of how molecular testing can affect the diagnostic workup for nodules of indeterminate cytology. Finally, we considered both test performance (sensitivity/specificity) and clinical utility (the probability of surgical resection after molecular testing) in our analysis, taking into consideration how the test results might affect clinical decision-making, which was more realistic and conservative.

Our analysis also had several limitations. First, our Markov model was simplified to capture the main benefits of molecular testing, such as reducing unnecessary surgeries. However, we did not have enough

data to model other aspects of the disease, such as cancer progression in missed malignant nodules or recurrence after surgery. Second, the prevalence of malignancy is a key model parameter that affects the cost-effectiveness of molecular testing; we used a point estimate from the literature (25%), but the prevalence of malignancy may vary greatly from institution to institution (e.g., high- vs. low-volume centres, different pathology practices). Our reference case results represent an average estimate for the entire population and may not be generalizable to all settings. In general, molecular testing would be more cost-effective when the prevalence of malignancy was lower. Third, we relied on values from the literature for health state utilities, and there were methodological issues with how these were generated, as mentioned above. More robust health state utility estimates may yield different results. Finally, the test sensitivities and specificities for Afirma GSC and ThyroSeq v3 were based on two clinical validation studies sponsored by the parent companies. The real-world performance of these tests may be different; more independent validation studies are needed.

Conclusions

Our primary economic evaluation found that molecular testing was more costly, but it increased the probability of predicting a correct diagnosis, reduced the probability of unnecessary surgery, and led to a slight improvement in QALYs compared to diagnostic lobectomy. The resulting ICERs were $220,572 to $298,653 per QALY gained, $4,451 to $6,328 per additional correct diagnosis, and $4,314 to $5,769 per unnecessary surgery avoided. At the commonly used willingness-to-pay values of $50,000 and $100,000 per QALY gained, molecular testing was moderately likely not to be cost-effective at its current list price.

Budget Impact Analysis

Research Question

What is the potential 5-year budget impact for the Ontario Ministry of Health of publicly funding molecular testing for people with thyroid nodules of indeterminate cytology?

Methods

Analytic Framework

We estimated the budget impact of publicly funding molecular testing for nodules of indeterminate cytology using the cost difference between two scenarios: (1) current clinical practice without public funding for molecular testing (the current scenario) and (2) anticipated clinical practice with public funding for molecular testing (the new scenario). Figure 8 presents the budget impact model schematic.

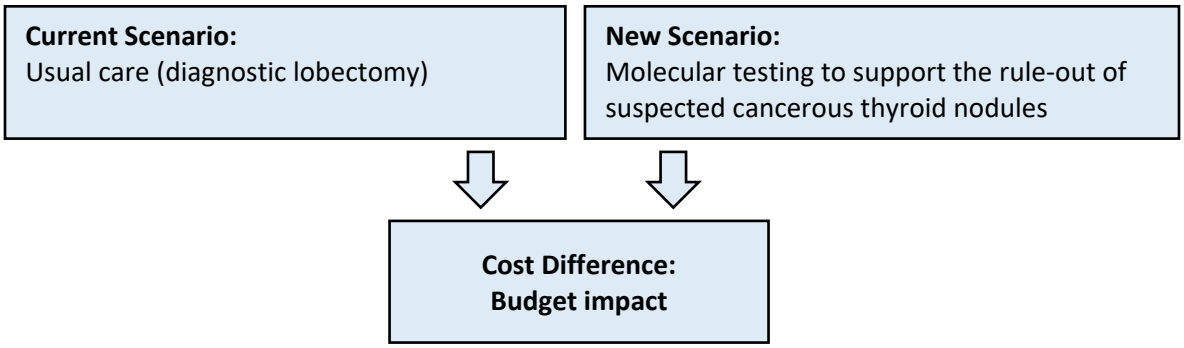

Figure 8: Schematic Model of Budget Impact

Key Assumptions

In addition to the assumptions made for the primary economic evaluation, we assumed the following for the budget impact analysis:

- The number of adults who had thyroid fine-needle aspiration biopsies each year would remain stable
- If molecular testing were publicly funded, the market shares of Afirma GSC and ThyroSeq v3 would be approximately equal

Target Population

Our target population was people who had a thyroid nodule of indeterminate cytology (The Bethesda System for Reporting of Thyroid Cytology [TBSRTC] III and IV). We estimated the current size of the target population based on epidemiology and administrative data (Figure 9).

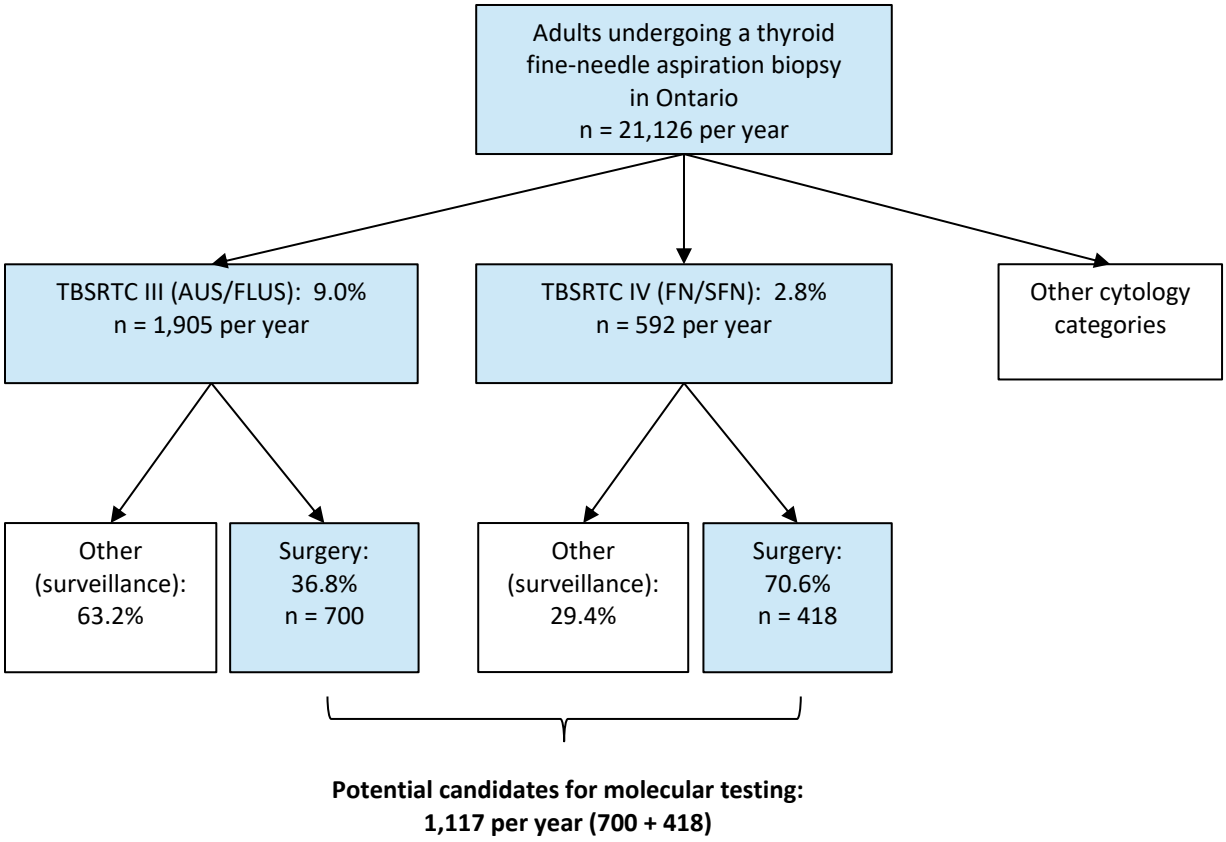

Figure 9: Flowchart to Estimate the Size of the Target Population

Abbreviations: AUS, atypical cell of undetermined significance; FLUS, follicular lesion of undetermined significance; FN, follicular neoplasm; SFN, suspicion for a follicular neoplasm; TBSRTC, The Bethesda System for Reporting Thyroid Cytology.
Note: Values may appear inexact due to rounding.
Source: Data prepared in April 2021 by Ontario Health using datasets from ICES, the Canadian Institute for Health Information, and the Ontario Ministry of Health.

First, we estimated the number of adults undergoing thyroid fine-needle aspiration biopsies using administrative data from the Ontario Health Insurance Plan (OHIP) claims database (Table 15). From 2016 to 2019 (the most recent 4 years with complete data), an average of 21,126 adults had at least one thyroid fine-needle aspiration biopsy per year (identified using OHIP code Z771). We assumed that the number of adults undergoing thyroid fine-needle aspiration biopsies would remain stable in subsequent years.

Table 15: Adults Who Had Thyroid Fine-Needle Aspiration Biopsies in Ontario

	2016	2017	2018	2019	Average
Number of adults with cytology claims in Ontario	22,769	21,148	19,635	20,952	21,126

Source: Ontario Health.

Next, we estimated the number of adults with a nodule of indeterminate cytology (TBSRTC III or IV). According to Ontario administrative data from 2016 to 2019, the average proportion of patients with TBSRTC III cytology is 9.0%, and the proportion of patients with TBSRTC IV cytology is 2.8%. These percentages have remained stable over the past 4 years. The proportions of patients with TBSRTC III and IV cytology results in Ontario were slightly lower than the values in the literature. According to a study by Bongiovanni et al,[16] a meta-analysis of 25,445 thyroid fine-needle aspiration samples reported from 8 studies using TBSRTC, 9.6% of all samples were TBSRTC III and 10.1% were TBSRTC IV.

Finally, we estimated the number of adults with a nodule of indeterminate cytology who might be suitable for molecular testing (Table 16). Because molecular tests are expensive (as discussed in the Primary Economic Evaluation), they should be used primarily to help physicians decide if a patient needs diagnostic lobectomy (surgery). Therefore, people who are considered for surgery would be good candidates for molecular testing. According to the Ontario administrative data from 2019, about 36.8% of patients with a nodule with TBSRTC III cytology and about 70.6% of patients with a nodule of TBSRTC IV cytology have undergone surgery. These Ontario numbers are very similar to the values in the literature. According to Bongiovanni et al,[16] about 39.2% of those with a nodule of TBSRTC III cytology and 69.7% of those with a nodule of TBSRTC IV cytology underwent surgery. However, it is possible that molecular testing may be used more widely among patients with nodules of indeterminate cytology. In a scenario analysis, we assumed that all patients with nodules of indeterminate cytology would be eligible for molecular testing.

Table 16: Target Population

	Year 1	Year 2	Year 3	Year 4	Year 5
Target population/volume, n	1,117	1,117	1,117	1,117	1,117
TBSRTC III	700	700	700	700	700
TBSRTC IV	418	418	418	418	418

Abbreviation: TBSRTC, The Bethesda System for Reporting Thyroid Cytology.
Note: Values may appear inexact due to rounding.

Current Intervention Mix

At present, molecular testing for thyroid nodules of indeterminate cytology is not publicly funded in Ontario. Therefore, we assumed that all patients in the current scenario were receiving usual care. In the reference case analysis, we assumed that all patients received diagnostic lobectomy as usual care. In a scenario analysis where we assumed that all individuals with nodules of indeterminate cytology were eligible for molecular testing, usual care consisted of both surveillance and diagnostic lobectomy.

Uptake of the New Intervention

We estimated the uptake rates of molecular testing in the new scenario based on consultation with various stakeholders (e.g., clinical experts, manufacturers, Ontario Ministry of Health; Table 17). According to clinical experts, if molecular tests were publicly funded, they would be adopted quickly, especially at high-volume centres (Antoine Eskander, email communication, March 28, 2021; Michael Odell, email communication, April 5, 2021). Our estimated rates were as follows:

- Reference case: 25%, 35%, 45%, 55%, and 65% in years 1 to 5
- Slow uptake scenario: 10% in year 1 to 50% in year 5
- Rapid uptake scenario: 80%, 85%, 90%, 95%, and 100% in years 1 to 5

Table 17: Uptake Rate of Molecular Testing in Ontario

Uptake	Year 1	Year 2	Year 3	Year 4	Year 5
Overall uptake rate, %	25	35	45	55	65
Afirma GSC	12.5	17.5	22.5	27.5	32.5
ThyroSeq v3	12.5	17.5	22.5	27.5	32.5
Overall volume, n	279	391	503	615	726
Afirma GSC	140	196	251	307	363
ThyroSeq v3	140	196	251	307	363

Abbreviation: GSC, gene sequencing classifier.
Note: Values may appear inexact due to rounding.

Resources and Costs

We estimated the annual per-person costs (undiscounted) associated with molecular testing strategy and usual care from the primary economic evaluation. The costs included those of molecular testing, surgery, management of surgical complications, ongoing surveillance, and hormone replacement therapy. All costs are reported in 2021 Canadian dollars.

Internal Validation

The secondary health economist conducted formal internal validation. This process included checking for errors and ensuring the accuracy of parameter inputs and equations in the budget impact analysis.

Analysis

We conducted a reference case analysis and sensitivity analyses. Our reference case analysis represented the analysis with the most likely set of input parameters and model assumptions. Our sensitivity analyses explored how the results were affected by varying input parameters and model assumptions.

We conducted sensitivity analyses by varying the following:

- Size of the target population (e.g., all patients with nodules of indeterminate cytology may be eligible for molecular testing)
- Prevalence (pre-test probability) of malignancy in the Ontario target population (6% and 46%)
- Assuming a second fine-needle aspiration biopsy is not needed
- Cost of molecular testing (± 25% to the list price)
- Uptake of molecular testing in years 1 to 5 (slow and rapid)

Results

Reference Case

The reference case results are presented in Table 18. In the current scenario, the cost of diagnostic lobectomy was about $6.53 million per year, for a total of $32.66 million over 5 years. In the new scenario, the cost of diagnostic lobectomy would decrease each year as the uptake of molecular testing increased, and the 5-year total cost would be about $17.96 million. The total health care cost for molecular testing would range from $2.33 million in year 1 to $6.05 million in year 5, for a total of $20.93 million over 5 years. We estimated the annual budget impact to be an additional $0.69 million in year 1, up to $1.80 million in year 5, for a total of $6.24 million over 5 years. We estimated the cost of providing molecular testing alone to be $1.34 million in year 1 to $3.48 million in year 5, for a total cost of $12.03 million over 5 years.

Table 18: Budget Impact Analysis Results—Reference Case

	Budget impact, $ in millions					
Scenario	**Year 1**	**Year 2**	**Year 3**	**Year 4**	**Year 5**	**Total**
Current scenario						
Diagnostic lobectomy	6.53	6.53	6.53	6.53	6.53	32.66
Cost of molecular test	0.00	0.00	0.00	0.00	0.00	0.00
Cost of fine-needle aspiration and additional physician visit	0.00	0.00	0.00	0.00	0.00	0.00
Cost of surgery and pathology	6.06	6.06	6.06	6.06	6.06	30.32
Cost of surveillance or follow-up	0.23	0.23	0.23	0.23	0.23	1.17
Cost of surgical complications and sequelae	0.23	0.23	0.23	0.23	0.23	1.17
New scenario						
Diagnostic lobectomy	4.90	4.25	3.59	2.94	2.29	17.96
Cost of molecular test	0.00	0.00	0.00	0.00	0.00	0.00
Cost of fine-needle aspiration and additional physician visit	0.00	0.00	0.00	0.00	0.00	0.00
Cost of surgery and pathology	4.55	3.94	3.34	2.73	2.12	16.68
Cost of surveillance or follow-up	0.18	0.15	0.13	0.11	0.08	0.64
Cost of surgical complications and sequelae	0.18	0.15	0.13	0.11	0.08	0.65
Molecular testing: Afirma GSC + ThyroSeq v3	2.33	3.26	4.19	5.12	6.05	20.93
Cost of molecular test	1.34	1.87	2.41	2.94	3.48	12.03
Cost of fine-needle aspiration and additional physician visit	0.06	0.08	0.10	0.13	0.15	0.52
Cost of surgery and pathology	0.73	1.02	1.32	1.61	1.90	6.58
Cost of surveillance or follow-up	0.17	0.24	0.31	0.38	0.45	1.56
Cost of surgical complications and sequelae	0.03	0.04	0.05	0.06	0.07	0.24
Budget impact	0.69	0.97	1.25	1.52	1.80	6.24
Cost of molecular testing	1.34	1.87	2.41	2.94	3.48	12.03

Abbreviation: GSC, gene sequencing classifier.

Sensitivity Analysis

The sensitivity analysis results are presented in Table 19. When we assumed a larger target population, the budget impact increased from $6.24 million to $13.91 million (scenario 1). When we assumed a lower prevalence of malignancy (scenario 2a) in the target population, the budget impact decreased. When we assumed a higher prevalence, the budget impact increased (scenario 2b). If molecular testing did not require an extra fine-needle aspiration biopsy to obtain tissue samples (i.e., the initial fine-needle aspiration biopsy sample could be used), the budget impact decreased slightly. Similarly, when we assumed that the cost of molecular testing was 25% lower, the budget impact also decreased. When we assumed that uptake rates were higher, the budget impact increased.

Table 19: Budget Impact Analysis Results—Sensitivity Analysis

	Budget impact, $ in millions					
Scenario	**Year 1**	**Year 2**	**Year 3**	**Year 4**	**Year 5**	**Total**
Reference case						
Budget impact	0.69	0.97	1.25	1.52	1.80	6.24
Cost of molecular testing	1.34	1.87	2.41	2.94	3.48	12.03
1: Size of the target population						
Budget impact	1.55	2.16	2.78	3.40	4.02	13.91
Cost of molecular testing	2.98	4.17	5.37	6.56	7.75	26.84
2a: Prevalence of malignancy, 6%						
Budget impact	0.54	0.76	0.98	1.20	1.42	4.90
Cost of molecular testing	1.34	1.87	2.41	2.94	3.48	12.03
2b: Prevalence of malignancy, 46%						
Budget impact	0.86	1.20	1.54	1.88	2.23	7.71
Cost of molecular testing	1.34	1.87	2.41	2.94	3.48	12.03
3: No extra fine-needle aspiration biopsy needed						
Budget impact	0.65	0.91	1.17	1.42	1.68	5.83
Cost of molecular testing	1.34	1.87	2.41	2.94	3.48	12.03
4a: Cost of molecular testing, +25%						
Budget impact	1.03	1.44	1.85	2.26	2.67	9.24
Cost of molecular testing	1.67	2.34	3.01	3.68	4.34	15.04
4b: Cost of molecular testing, −25%						
Budget impact	0.36	0.50	0.65	0.79	0.93	3.23
Cost of molecular testing	1.00	1.40	1.80	2.21	2.61	9.02
5a: Slow uptake						
Budget impact	0.28	0.55	0.83	1.11	1.39	4.16
Cost of molecular testing	0.53	1.07	1.60	2.14	2.67	8.02
5b: Rapid uptake						
Budget impact	2.22	2.36	2.49	2.63	2.77	12.47
Cost of molecular testing	4.28	4.54	4.81	5.08	5.35	24.06

Abbreviation: GSC, gene sequencing classifier.

Discussion

Publicly funding molecular testing in people with thyroid nodules of indeterminate cytology may add additional costs to the provincial budget (about $6.24 million over the next 5 years, for about 1,117 people each year). The costs of providing the test itself would be much higher (about $12 million over the next 5 years). Because molecular tests are expensive ($4,785/test), it would be important to identify clearly who should be eligible for these tests in Ontario. According to the recent INESSS report,[18] "the use of molecular tests should be limited to those whose test result is likely to influence management. Molecular testing should not be proposed to those for whom surgery is indicated (high-risk factors), those for whom diagnostic lobectomy is not being considered because of a low risk of malignancy, or those who express a preference for surgery or observation." To help manage costs, there could be some agreement with the manufacturers to share the financial risks.

Strengths and Limitations

Our analysis had several strengths. We estimated the size of the target population using real-world data from Ontario administrative databases. We considered not only the cost of the molecular test but also the potential cost consequences related to surgeries, management of surgical complications and sequelae, and ongoing surveillance. We also estimated the budget impact of several different scenarios by varying the prevalence of malignancy, testing costs, and uptake rates. A limitation is that we did not estimate the costs related to implementation, service delivery, or program coordination, because these could vary substantially depending on how testing is implemented.

Conclusions

Publicly funding molecular testing for thyroid nodules of indeterminate cytology may increase the budget by about $0.69 million in year 1, up to $1.80 million in year 5, for a total of $6.24 million over the next 5 years. We estimate the additional cost required for the molecular tests alone to be about $1.34 million in year 1, up to $3.48 million in year 5, for a total of $12.03 million over the next 5 years.

Preferences and Values Evidence

Objective

The objective of this analysis was to explore the underlying values, needs, and priorities of those who had lived experience of thyroid nodules and had considered (or undergone) molecular testing for nodules of indeterminate cytology. We also sought to understand patients' perceptions of risk and decision-making when they were choosing whether or not to undergo molecular testing and subsequent treatment.

Background

Exploring patient preferences and values provides a unique source of information about people's experiences of a health condition and the health technologies or interventions used to manage or treat that health condition. It includes the impact of the condition and its treatment on the person with the health condition, their family and other caregivers, and the person's personal environment. Engagement also provides insights into how a health condition is managed by the province's health system.

Information shared from lived experience can also identify gaps or limitations in published research (e.g., outcomes important to those with lived experience that are not reflected in the literature).[114-116] Additionally, lived experience can provide information and perspectives on the ethical and social values implications of health technologies or interventions.

Because the needs, preferences, priorities, and values of those with lived experience in Ontario are important to consider, to understand the impact of the technology in people's lives, we may speak directly with people who live with a given health condition, including those with experience of the technology or intervention we are exploring.

For this analysis, we examined the preferences and values of people with thyroid nodules of indeterminate cytology who underwent molecular testing in two ways:

- A review by Ontario Health of the quantitative evidence on patient preferences and values
- Direct engagement by Ontario Health with people who had experience with molecular testing for thyroid nodules—or who may encounter molecular testing—through interviews

Quantitative Evidence

Building on the literature search and screening conducted for the clinical evidence review in this health technology assessment, we leveraged the report published by INESSS.[18] Because it was developed from the perspective of another Canadian province, this report was the most contextually relevant, and it had up-to-date evidence (including the most recently published available evidence). We have summarized the key findings of this report below.

Key Findings

- The INESSS[18] report examined the care experience of people seeking a diagnosis for a suspicious thyroid nodule. It also explored concerns about molecular testing among people with findings of indeterminate cytology for a thyroid nodule. Two studies were included that addressed these questions as they related to the patient perspective.
- Lee et al[117] posed hypothetical scenarios to 100 people who were referred to an otolaryngology clinic in Toronto, half of whom had a thyroid nodule of indeterminate cytology.
- Wong et al[118] surveyed 332 people who had received cytology results and The Bethesda System for Reporting Thyroid Cytopathology (TBSRTC) classification, including 58 people who had an indeterminate classification (TBSRTC III or IV) and had received a molecular assessment using the Afirma GEC or ThyroSeq v2.1. Of these people, 37 had nodules that were molecularly benign.
- The findings of the INESSS[18] report suggested that molecular testing may provide reassurance of a nodule's benign status for those with thyroid nodules of indeterminate cytology. People may prefer molecular testing to standard care alone, but this choice is influenced by cost and not by the accuracy of the test.

Direct Patient Engagement

Methods

PARTNERSHIP PLAN

The partnership plan for this health technology assessment focused on consultation to examine the experiences of people with thyroid nodules and molecular testing. We engaged people via phone interviews.

We used a qualitative interview, as this method of engagement allowed us to explore the meaning of central themes in the experiences of people with thyroid nodules.[119] The sensitive nature of exploring people's experiences of a health condition and their quality of life were other factors that supported our choice of an interview methodology.

PARTICIPANT OUTREACH

We used an approach called purposive sampling,[120-123] which involves actively reaching out to people with direct experience of the health condition and health technology or intervention being reviewed. We approached a variety of partner organizations and clinical experts to spread the word about this engagement activity and to contact people who had experience with thyroid nodules and molecular testing.

Inclusion Criteria

We sought to speak with people who had lived experience of thyroid nodules and thyroid molecular testing, or those who may seek out molecular testing in the future. Participants did not need to have direct experience with thyroid molecular testing.

Exclusion Criteria

We did not set specific exclusion criteria.

Participants

For this project, we spoke to 13 people who had diagnosed thyroid nodules, living in Ontario. We spoke to people whose thyroid nodules were benign, indeterminate, or malignant. Participants were mostly from southern Ontario, but we also spoke to people from Thunder Bay, North Bay, and the Kingston area.

APPROACH

At the beginning of the interview, we explained the role of our organization, the purpose of this health technology assessment, the risks of participation, and how participants' personal health information would be protected. We gave this information to participants both verbally and in a letter of information (Appendix 7) if requested. We then obtained participants' verbal consent before starting the interview. With participants' consent, we audio-recorded and then transcribed the interviews.

Interviews lasted approximately 15 to 30 minutes. The interview was loosely structured and consisted of a series of open-ended questions. Questions were based on a list developed by the Health Technology Assessment International Interest Group on Patient and Citizen Involvement in Health Technology Assessment.[124] Questions focused on the diagnosis of thyroid nodules, participants' care journey and their perceptions of thyroid molecular testing and the ultimate impact of their treatment. See Appendix 8 for our interview guide.

DATA EXTRACTION AND ANALYSIS

We used a modified version of a grounded-theory methodology to analyze interview transcripts. The grounded-theory approach allowed us to organize and compare information on experiences across participants. This method consists of a repetitive process of obtaining, documenting, and analyzing responses while simultaneously collecting, analyzing, and comparing information.[125,126] We used the qualitative data analysis software program NVivo[127] to identify and interpret patterns in the data. The patterns we identified allowed us to highlight the impact of thyroid nodules and treatments, as well as the decision-making process for thyroid molecular testing from those we interviewed.

Results

DIAGNOSIS OF THYROID NODULES

The people we interviewed said that they were generally unaware that they had thyroid nodules until their physician discovered them. The nodules were most often discovered under ultrasound and were often found incidentally, in the course of other medical investigations. A few participants mentioned that blood work had revealed abnormal thyroid hormone levels, and the nodules were discovered as a result; others reported that their nodules were found as part of unrelated investigations into neck pain, tooth pain, or other ailments:

> *I don't know what prompted the ultrasound. I can't remember. It was back in 2001 that I had an ultrasound of the throat by my GP [general practitioner], and at the time when the results were received, they were identified as nodules on my thyroid. I guess [an ultrasound] was prompted by a blood test; maybe my thyroid levels were high. I guess they would be high.*

> *I was in the ICU [intensive care unit]. I had a hypernatremia episode. I was unconscious, and then while I was unconscious in order to figure out what was wrong, they did a whole bunch of different scanning. In the process of those scans, they found two masses. One was this thyroid nodule.*

> *I had swollen lymph nodes. They felt really big, and I was concerned about them. So my doctor sent me for an ultrasound and actually while they were checking my lymph nodes, they found the cysts … the nodules on my thyroid.*

We asked participants if symptoms or other signals had indicated the presence of thyroid nodules. Some reported experiencing minor symptoms, such as a small amount of neck swelling, uncomfortable swallowing, or sensations of fatigue that they associated with the nodules, but these did not point to anything conclusive. Most participants reported no remarkable symptoms at all, and no previous hint that nodules had developed or when they developed. Because of this, the presence of the nodules did not necessarily affect participants' day-to-day activities or quality of life:

> *I don't remember exactly how they came about. But I think we did a blood test back in 2000, and I think there had been some symptoms, but the blood test showed that there was some issue with the thyroid.*

> *All of a sudden, I felt like I had something stuck in my throat. You know, when you have something to eat, you got a little piece of something. But it just didn't go away. So then that's when I went to the doctor and did an ultrasound and then went to see an ENT [ear, nose, and throat] specialist to look down there. And [he] said, "Yeah, there is a growth …"*

Despite the lack of worrying symptoms from the thyroid nodules or any significant effect on day-to-day activities, participants did report that the discovery of the nodules led to a variety of emotions. Some participants felt reassured by their physicians that the nodules were most likely benign and reported feeling little anxiety. Others were more anxious about the nature of the nodules and their potential to grow or become cancerous. They spoke about their fears, their sense of dread about the potential for malignancy, and how their feelings changed over time:

> *Well, I guess I was very concerned, because I didn't know what it was—if it was cancer. And I didn't realize it was that big.*

> *I mean, it was a surprise. "Oh, isn't that interesting." … I thought, is that why I sometimes have trouble swallowing or various other things that are in that area?*

> *At that time, for sure, I thought, oh, that's great … I'm figuring [the doctors] see a lot so can tell different variations of [cancer versus noncancer growth]. But then as you get a little bit older, you think a little bit more, maybe it's something … maybe there is more that needs to be looked at.*

A couple of the people we spoke to had a history of thyroid issues in their family, including cancer. For these participants, the familial experience meant that their level of anxiety and fear were high, and they were concerned about the presence of thyroid nodules:

> *I was scared because my cousin who was younger than me had thyroid cancer. She had a really big [nodule]—like it was huge. And she had that removed in her 20s. So I was concerned because of that.*

CARE JOURNEY

After their thyroid nodules had been discovered, participants reported having two basic options when it came to pathways of care: watchful waiting or further investigation.

Those who were told that their nodules were benign reported a level of comfort with ongoing regular ultrasounds to monitor the nodules and check to see if they had grown or changed. Some of these participants had not even undergone a biopsy, because their physicians deemed it unnecessary, and others expressed reservations about having a biopsy done at all if it was not urgently needed:

> *So anyway, then they said, "Well, you know usually this is benign. Let's just watch it." I said, "Watchful waiting, I'm all for that."*

> *Yeah, [a biopsy] hasn't come up. So I'm just presuming … there is no problem. So it's not something that's been brought to me, and I haven't brought that up, either.*

> *[A biopsy] sounded a pretty invasive to me. I hate needles. And my naturopath and I did discuss this, and we're into watchful waiting. You know, it doesn't freak me out that the nodule is there.*

Other participants did have a biopsy done, the results of which could indicate that the nodule was benign, cancerous, or indeterminate. Some participants reported having multiple biopsies done over a number of years and the anxiety that accompanied the wait for results:

> *And at that point the doctor had done a biopsy, and [it was] 5 weeks for the biopsy to come back. Yeah, like nerves of steel.*

> *I think I've had it done three times now, and that's okay. Yeah, about every five years they redo it just in case.*

Naturally, some participants reported that there had been changes in their care pathway, including new physicians wanting further investigations, or new concerns about the growth or spread of the nodules. Some participants felt as if the nodule had changed and pushed for further imaging. This prompted participants to move from watchful waiting to more active investigation:

> *They said that I had nodules, but they didn't think I needed treatment. So nothing happened for about 3 years. And I changed doctors, and the new doctor took a look and found this and had a fit because I hadn't told him about it, and I hadn't told him because I didn't know it was an issue. Yeah, so he took a look at it and said, "Well, it's definitely an issue." Did an ultrasound and found out the nodules were growing.*

> *Sometimes it's me going in and saying, "Look, I feel there's been a change, I feel like it's bigger." And then other times [the doctor will say], "Okay, let's check it again."*

Several participants reported circumstances that would prompt them to actively switch from watchful waiting to more active investigations, including physical changes to the nodule, worrying new symptoms, or the urging of their physicians:

> *I think it would firstly start with how I'm feeling; if I noticed any big changes, then I would go and confer further about that. And then obviously, if I had a follow-up ultrasound or something and*

they're saying, "Oh yeah, there's been some growth or some change," then that would lead me more to say, "Yes definitely, I'd like to get that checked."

So my take on this would be, if I'm getting more symptoms—say my thyroid starts hurting or swelling, or the next test said those nodules are growing and they were interfering with swallowing and whatnot, seriously. Well, then I'd be interested in possibly a biopsy and maybe the genetic test if that would be useful for sure. I'm not at that point right now … and I just had a birthday yesterday, so I'm hoping that I'll outlast this whole deal.

INDETERMINATE CYTOLOGY AND DECISION-MAKING

Participants who reported that their biopsy results were indeterminate had several options for how to proceed. Participants could continue to monitor their nodules, redo the biopsy, undergo molecular testing to further refine the results, or opt for surgery:

And when it did come back it showed that it was inconclusive. So when I heard "inconclusive" I thought, "Hmm, that doesn't sound like a good thing." And the three options I was given were … to rebiopsy again in 6 months, to sign on for this molecular testing, which was not available in Canada, and to have it surgically removed.

Making the decision about whether to proceed with surgery or not was a unique process for each person. Participants weighed multiple factors, such as the risks of surgery, the risks of potential cancer growth, their emotions about surgery, and long-term consequences. They often reviewed these considerations in consultation with their physician and family members, and they did their own research into their condition and potential effects of the surgery. Several participants carefully explained the details of their own process and preferences for choosing to undergo surgery or not:

Well, because I was over 50 years of age when I found it and the risks were more elevated, I was the one who asked for a total thyroidectomy. They sort of recommended it—[the doctor] was not sure, but sort of recommended that it would be just a unilateral thyroidectomy, just half of it. And after doing a bit of research, I felt that I would feel better if I went through the whole thing instead of going back for a second surgery. And sure enough, if I had only the partial, all of these 3½ months, it would have been a long time not knowing if I needed the second surgery.

For some, historical or familial factors meant that they leaned toward having surgery or away from it. Often the experiences of friends or family members in similar circumstances were a factor in their decision-making:

The surgeon had said that between 20% and 30% of these cases are cancerous, but likely it's not. And those odds seemed a little bit high to me. Like I wasn't comfortable with 1 out of 5 cases or 1 out of 4 cases. And this particularly since one of my aunts had died of thyroid cancer many years ago. They didn't think that was a factor because it wasn't a direct relation, but in terms of where I was comfortable with, I opted for the surgery.

And he left it with me, and I sat on it for awhile. And then I checked back with my ENT [ear, nose, and throat] doctor and my GP [general practitioner] and my son who's a doctor and a couple of friends who have gone through not thyroid issues but other surprises in their lives. And they all said, "Well why do you want to have to worry about it? Just have it removed."

Another factor that seemed to influence participants was the potential need for supplementary medications long-term if a full thyroidectomy was performed. This consequence was concerning for some participants, but not for others:

> *I guess the certainty of surgery [was a factor]. I probably thought to myself, "Well, if he takes it out, I'm certainly going to have to take meds for the rest of my life."*

> *Yeah, I guess if you don't have that [family] history and you aren't aware of things you might jump at maybe going to get the surgery. I think it might seem scarier, I guess. I've just seen how my mom and my grandma have managed it with taking their thyroid pills.*

These sentiments about medications are one example of the personal preferences that went into decision-making around surgery. A few participants commented explicitly on the personal nature of the decision: how personality may influence decision outcomes. Some participants may simply want the risk of cancer to be removed, so surgery was a much more feasible and less stressful option:

> *I think somewhat it's reflective of people's individual personalities. And I don't know if I can also generalize to sexes, but at least within us as a couple, my husband is of the mindset if there's something there, cut it out, get rid of it. So when someone suggested that you might want to have surgery to be cautious, he probably said fine, go ahead. Where I would I probe things to a deeper level than he did. So that's part of it; those personality differences that lead to different results.*

Naturally, these important decisions and the potential serious consequences could be burdensome. A number of participants reflected on their levels of fear and anxiety at different stages of the process and the emotional toll it took on their lives. In contrast, a couple of participants who were facing other serious health concerns viewed thyroid issues with less anxiety or concern:

> *You're living on pins and needles and wondering whether you're going to need [surgery] … it was a very, very long time, those 3½ months. It was [stressful] for me and my family.*

> *I feel like this uncertainty bothers me a lot, and that's probably why I have asked for another opinion. I don't feel like I know enough about what they're looking for other than cancer. So the specialist also said that I don't require surgery, but I don't feel comfortable. I always feel like maybe they're not looking hard enough, maybe I should be monitored more closely than "next ultrasound is in 3 months." So I have this uneasy feeling about that.*

> *If this just arose in my normal life and I hadn't just had all of these episodes I would definitely have a lot more anxiety about it. Because the idea of having a cancer in general is a very scary thing. And surgery in general is a scary thing, 100%, but coming from where I'm coming from on the relativism scale. I'm like, "Well, 99.5% of cases of thyroid cancer are very curable."*

MOLECULAR TESTING: NODULES OF INDETERMINATE CYTOLOGY

Most of the people we interviewed were unaware of molecular testing for nodules of indeterminate cytology. During the interview, we presented information about molecular testing for nodules of indeterminate cytology to all participants, even if their nodules were benign, and asked them to consider its implications. We asked them about the hypothetical effect molecular testing could have had on their decision-making, and about any benefits or drawbacks they saw in using the test.

Almost all participants reported positive impressions of the potential usefulness of the test and its ability to provide more information for decision-making. They saw it as a way to refine decision-making and as another tool for both the physician and patient to use when deciding about surgery. However, they connected the usefulness of molecular testing to the relative accuracy of the test and its ability to provide a definitive answer. Some participants mentioned that if the test result was uncertain and provided no new definitive information, then time would likely have been wasted, perhaps allowing cancer to spread further:

> *No, and I wouldn't want to do it … I would not want to do surgery unless I had to. So if this test would help me make that decision, I would definitely try it.*

> *I think I would be willing to do the molecular testing, wait 6 weeks and then proceed with surgery if I needed to. But that would depend on the understanding of the efficacy of the molecular testing. If I would know with 90% to 95% … If I would know with enough certainty for someone to make a decision based on the molecular testing, then I would say yes. But if it was going to be another 6 weeks and then it was still going to have to be some kind of judgment call or speculation and I was going to need to be surgery anyway, then I probably wouldn't want to delay it, if that makes sense.*

> *I would want to know how reliable that test would be. And also if there's pain involved—always ask that question. … Those are my main concerns: how reliable it is and also pain. And then timing; the one that I did took months to come back.*

A few participants also mentioned an overall hesitancy with genetic tests in general, wanting to be assured that the data provided were stored safely and could not be lost. However, this was not a majority opinion; most others were less concerned about this aspect of molecular testing:

> *I would say not as a medical professional but as an informed person, what would make me concerned about genetic testing right now is what is happening to the data.*

> *Well, as long as the information is well protected, you know, any genetic information is very highly easy to steal. That would be my one concern, but otherwise more information would have been very helpful in deciding whether I needed a total or partial thyroidectomy.*

> *Well now that you mention it, I've had one genetic test done before, and you just kind of like put your saliva on this thing, and [it], goes out, you know, and then comes back as a report. So, if it was anything like that, I wouldn't necessarily be afraid of it.*

A number of participants with nodules of indeterminate cytology spoke about how they had considered using the molecular test, but that they had multiple reasons whey they did not choose to. A common reason was the delay in obtaining test results; a second biopsy would need to be done to obtain a sample for the molecular test, and then the test would be sent off for analysis. This process could add several months of waiting, and most participants mentioned that this was not something they wanted to do. Participants expressed fear about the potential for a malignancy to grow during that time delay:

> *And the reason I didn't opt for the molecular test is that I would have had to wait another 6 months for the next biopsy and then another potentially 5 weeks to get results, and I thought, "That's a lot of time wasted if in fact there's something looming in there." That's my story.*

> *Yeah, the cost was not the issue for me. It was the time factor, because it would have extended getting any results by at least 7 months. And I thought by the time I get the results if they've fallen into that 20% to 30% and then have to book surgery, I've lost about 8 months. And some thyroid cancers are very, very aggressive, and that's what my worry was. So I thought for peace of mind …*

Other participants acknowledged that the cost of the molecular test may be a factor in choosing it or not. Those who had considered using the molecular test were able to access it through a research program, but they acknowledged that the full cost of the molecular test may be a barrier and could influence decision-making, depending on the cost point:

> *I guess the first would be if the cost was exorbitant, then I would probably be less inclined to do it if I didn't know if there would be more efficacy. But my understanding is it was sort of like in the $3,000 to $5,000 range, which makes it something that … if that would decrease your likelihood of needing more surgery based on just conservativism, that's probably something that I would do. I'm just saying, if it was like $100,000, I think I would be like, yeah, I'll just get surgery at that point, I think.*
>
> *Yeah, and I can't afford genetic tests, I'm on a very small pension. I'm just barely managing here.*

BARRIERS

Beyond the costs associated with molecular testing, some participants mentioned other barriers in accessing care, including time delays. Because of the fear of potential of growth in a malignancy, any delay in care caused anxiety for participants. Some participants spoke of multiple delays in diagnosis and treatment, which caused emotional distress. Possible delays also included finding available specialists to perform biopsies or review results. This was particularly true of participants who lived in Northern Ontario. Other participants felt that they were seen in a relatively timely manner and did not encounter any noticeable barriers:

> *Yeah, I think if I was given the choice, I would have wanted it done right away. So I did wait 2 months for the [surgery]. And that 2 months did seem like a long time, but in hindsight it wasn't that long.*
>
> *Nine months, not just to get the new results, but to wait for the biopsy, to extract a sample, to then send off. It was just a lot of time wasted.*
>
> *I think it's pretty good. Obviously, to get to a specialist of any kind it takes a couple months, right? If I was to say to the doctor, "Oh, I feel these changes. I'd like to go back and have that reassessed by that specialist," it would take I'm sure a couple months to get in in wait time, but not I would say like over a year or anything like that.*

Discussion

Through direct engagement, we performed a robust examination of patient preferences and values related to thyroid nodules and decision-making about molecular testing. All participants had had a diagnosis of thyroid nodules and direct experience with analysis and decision-making around potential treatments.

We were also able to speak with participants who had followed different care pathways for their thyroid nodules, from active surveillance to full thyroidectomies. This allowed for a thorough exploration of decision-making factors and patient preferences when considering their care. A number of participants had received a diagnosis of indeterminate cytology for their thyroid nodules, allowing us to ask directly about their decision-making related to the use of molecular testing for subsequent analysis. Participants spoke to the benefits and drawbacks of molecular testing, allowing for analysis of its potential impact.

Participants were also able to speak about barriers they may have faced in accessing and choosing molecular testing for nodules of indeterminate cytology. This context can provide insight into the use of molecular testing in Ontario and help to illuminate when and how participants may access the test.

A limitation of this engagement was the relatively limited number of participants who have used molecular testing for nodules of indeterminate cytology. Molecular testing is not widely available in Ontario and is only accessed through out-of-pocket payment or research studies.

Conclusions of Direct Patient Engagement

Participants reported anxiety and fear of malignancy as a result of the discovery of thyroid nodules. Although participants saw molecular testing as a potential opportunity to obtain further information and aid in treatment decision-making, they were concerned that the results could be delayed or that they would not conclusively aid in treatment decision-making.

Preferences and Values Evidence Discussion

We obtained patient preferences and values by examining the quantitative evidence and conducting direct patient engagement. Each method examined people's perspectives on molecular testing for nodules of indeterminate cytology.

The quantitative evidence leveraged a report by INESSS,[18] and its findings aligned with the results of the direct engagement: people had a strong perception of value of the molecular test in providing information to help with treatment decision-making. However, the usability of this test could be tied to cost, a sentiment expressed in both the INESSS report and through direct patient engagement. As well, the people we interviewed expressed concern that if the molecular test did not provide accurate and decisive information, it may be of less value.

Preferences and Values Evidence Conclusions

The discovery of thyroid nodules can cause people anxiety and concern about the potential for malignancy. In the literature and through direct engagement, we found that people valued molecular testing of thyroid nodules as an opportunity to obtain further information that could aid in treatment decision-making. Participants expressed concern about the time required to obtain the results of molecular testing, particularly if the findings were inconclusive or of little use in treatment decision-making.

Conclusions of the Health Technology Assessment

Molecular testing for thyroid nodules of indeterminate cytology has a sensitivity of 91% to 94% and a specificity of 68% to 82% for the detection of malignancy. As well, lower rates of surgical resections were reported in nodules of indeterminate cytology compared to usual (molecular testing vs. no molecular testing), but the evidence is very uncertain.

Our primary economic evaluation found that molecular testing was more costly, but it increased the probability of predicting a correct diagnosis, reduced the probability of unnecessary surgery, and led to a slight improvement in QALYs compared to diagnostic lobectomy. The resulting ICERs were $220,572 to $298,653 per QALY gained, $4,451 to $6,328 per additional correct diagnosis, and $4,314 to $5,769 per unnecessary surgery avoided. At the commonly used willingness-to-pay values of $50,000 and $100,000 per QALY gained, molecular testing was unlikely to be cost-effective at its current price.

Publicly funding molecular testing for thyroid nodules of indeterminate cytology may increase the budget by about $0.69 million in year 1, up to $1.80 million in year 5, for a total of $6.24 million over the next 5 years (at a cost of $4,785 per test for 1,117 people per year). We estimate the additional cost required for the molecular tests alone to be about $1.34 million in year 1, up to $3.48 million in year 5, for a total of $12.03 million over the next 5 years.

The discovery of thyroid nodules can cause people anxiety and concern about the potential for malignancy. In the literature and through direct engagement, we found that people valued molecular testing of thyroid nodules as an opportunity to obtain further information that could aid in treatment decision-making. Participants expressed concern about the time required to obtain the results of molecular testing, particularly if the findings were inconclusive or of little use in treatment decision-making.

Abbreviations

AUS	Atypical cell of undetermined significance
CADTH	Canadian Agency for Drugs and Technologies in Health
FLUS	Follicular lesion of undetermined significance
GEC	Gene expression classifier
GMP	Gene mutation panel
GRADE	Grading of Recommendations Assessment, Development, and Evaluation
GSC	Genomic sequence classifier
ICER	Incremental cost-effectiveness ratio
INESSS	Institut national d'excellence en santé et en services sociaux
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-analyses
QALY	Quality-adjusted life-year
QUADAS-2	Quality Assessment of Diagnostic Accuracy Studies 2
RoBANS	Risk of Bias tool for Non-randomized Studies
ROBIS	Risk of Bias in Systematic Review
TBSRTC	The Bethesda System for Reporting Thyroid Cytopathology

Glossary

Budget impact analysis	A budget impact analysis estimates the financial impact of adopting a new health care intervention on the current budget (i.e., the affordability of the new intervention). It is based on predictions of how changes in the intervention mix will impact the level of health care spending for a specific population. Budget impact analyses are typically conducted for a short-term period (e.g., 5 years). The budget impact, sometimes referred to as the net budget impact, is the estimated cost difference between the current scenario (i.e., the anticipated amount of spending for a specific population without using the new intervention) and the new scenario (i.e., the anticipated amount of spending for a specific population following the introduction of the new intervention).
Cost-effective	A health care intervention is considered cost-effective when it provides additional benefits, compared with relevant alternatives, at an additional cost that is acceptable to a decision-maker based on the maximum willingness-to-pay value.
Cost-effectiveness analysis	Used broadly, "cost-effectiveness analysis" may refer to an economic evaluation used to compare the benefits of two or more health care interventions with their costs. It may encompass several types of analysis (e.g., cost-effectiveness analysis, cost–utility analysis). Used more specifically, "cost-effectiveness analysis" may refer to a type of economic evaluation in which the main outcome measure is the incremental cost per natural unit of health (e.g., life-year, symptom-free day) gained.
Cost–utility analysis	A cost–utility analysis is a type of economic evaluation used to compare the benefits of two or more health care interventions with their costs. The benefits are measured using quality-adjusted life-years, which capture both the quality and quantity of life. In a cost–utility analysis, the main outcome measure is the incremental cost per quality-adjusted life-year gained.
Decision tree	A decision tree is a type of economic model used to assess the costs and benefits of two or more alternative health care interventions. Each intervention may be associated with different outcomes, which are represented by distinct branches in the tree. Each outcome may have a different probability of occurring and may lead to different costs and benefits.
Discounting	Discounting is a method used in economic evaluations to adjust for the differential timing of the costs incurred and the benefits generated by a health care intervention over time. Discounting reflects the concept of positive time preference, whereby future costs and benefits are reduced to reflect their present value. The health technology assessments conducted by Ontario Health use an annual discount rate of 1.5% for both future costs and future benefits.

Disutility	A disutility is a decrease in utility (i.e., a decrease in preference for a particular health outcome) typically resulting from a particular health condition (e.g., experiencing a symptom or complication).
Health-related quality of life	Health-related quality of life is a measure of the impact of a health care intervention on a person's health. It includes the dimensions of physiology, function, social life, cognition, emotions, sleep and rest, energy and vitality, health perception, and general life satisfaction.
Health state	A health state is a particular status of health (e.g., sick, well, dead). A health state is associated with some amount of benefit and may be associated with specific costs. Benefit is captured through individual or societal preferences for the time spent in each health state and is expressed in quality-adjusted weights called utility values. In a Markov model, a finite number of mutually exclusive health states are used to represent discrete states of health.
Incremental cost	The incremental cost is the additional cost, typically per person, of a health care intervention versus a comparator.
Incremental cost-effectiveness ratio (ICER)	The incremental cost-effectiveness ratio (ICER) is a summary measure that indicates, for a given health care intervention, how much more a health care consumer must pay to get an additional unit of benefit relative to an alternative intervention. It is obtained by dividing the incremental cost by the incremental effectiveness. Incremental cost-effectiveness ratios are typically presented as the cost per life-year gained or the cost per quality-adjusted life-year gained.
Markov model	A Markov model is a type of decision-analytic model used in economic evaluations to estimate the costs and health outcomes (e.g., quality-adjusted life-years gained) associated with using a particular health care intervention. Markov models are useful for clinical problems that involve events of interest that may recur over time (e.g., stroke). A Markov model consists of mutually exclusive, exhaustive health states. Patients remain in a given health state for a certain period of time before moving to another health state based on transition probabilities. The health states and events modelled may be associated with specific costs and health outcomes.
Ministry of Health perspective	The perspective adopted in economic evaluations determines the types of costs and health benefits to include. Ontario Health develops health technology assessment reports from the perspective of the Ontario Ministry of Health. This perspective includes all costs and health benefits attributable to the Ministry of Health, such as treatment costs (e.g., drugs, administration, monitoring, hospital stays) and costs associated with managing adverse events caused by treatments. This perspective does not include out-of-pocket costs incurred by patients related to obtaining care (e.g., transportation) or loss of productivity (e.g., absenteeism).
Natural history of a disease	The natural history of a disease is the progression of a disease over time in the absence of any health care intervention.

Probabilistic analysis	A probabilistic analysis (also known as a probabilistic sensitivity analysis) is used in economic models to explore uncertainty in several parameters simultaneously and is done using Monte Carlo simulation. Model inputs are defined as a distribution of possible values. In each iteration, model inputs are obtained by randomly sampling from each distribution, and a single estimate of cost and effectiveness is generated. This process is repeated many times (e.g., 10,000 times) to estimate the number of times (i.e., the probability) that the health care intervention of interest is cost-effective.
Quality-adjusted life-year (QALY)	The quality-adjusted life-year (QALY) is a generic health outcome measure commonly used in cost–utility analyses to reflect the quantity and quality of life-years lived. The life-years lived are adjusted for quality of life using individual or societal preferences (i.e., utility values) for being in a particular health state. One year of perfect health is represented by one quality-adjusted life-year.
Reference case	The reference case is a preferred set of methods and principles that provide the guidelines for economic evaluations. Its purpose is to standardize the approach of conducting and reporting economic evaluations, so that results can be compared across studies.
Scenario analysis	A scenario analysis is used to explore uncertainty in the results of an economic evaluation. It is done by observing the potential impact of different scenarios on the cost-effectiveness of a health care intervention. Scenario analyses include varying structural assumptions from the reference case.
Time horizon	In economic evaluations, the time horizon is the time frame over which costs and benefits are examined and calculated. The relevant time horizon is chosen based on the nature of the disease and health care intervention being assessed, as well as the purpose of the analysis. For instance, a lifetime horizon would be chosen to capture the long-term health and cost consequences over a patient's lifetime.
Time trade-off	In economic evaluations, time trade-off is a direct method of measuring people's preferences for various health states. In a time-trade off, respondents are asked about their preference for either (a) living with a chronic health condition for a certain amount of time, followed by death, or (b) living in optimal health but for less time than in scenario (a). That is, respondents decide how much time in good health they would be willing to "trade off" for more time spent in poorer health. Respondents are surveyed repeatedly, with the amount of time spent in optimal health varying each time until they are indifferent about their choice.
Uptake rate	In instances where two technologies are being compared, the uptake rate is the rate at which a new technology is adopted. When a new technology is adopted, it may be used in addition to an existing technology, or it may replace an existing technology.

Utility	A utility is a value that represents a person's preference for various health states. Typically, utility values are anchored at 0 (death) and 1 (perfect health). In some scoring systems, a negative utility value indicates a state of health valued as being worse than death. Utility values can be aggregated over time to derive quality-adjusted life-years, a common outcome measure in economic evaluations.
Willingness-to-pay value	A willingness-to-pay value is the monetary value a health care consumer is willing to pay for added health benefits. When conducting a cost–utility analysis, the willingness-to-pay value represents the cost a consumer is willing to pay for an additional quality-adjusted life-year. If the incremental cost-effectiveness ratio is less than the willingness-to-pay value, the health care intervention of interest is considered cost-effective. If the incremental cost-effectiveness ratio is more than the willingness-to-pay value, the intervention is considered not to be cost-effective.

Appendices

Appendix 1: Literature Search Strategies

Clinical Evidence Search

Search date: September 9, 2020

Databases searched: Ovid MEDLINE, Embase, Cochrane Central Register of Controlled Trials, Cochrane Database of Systematic Reviews, CRD Health Technology Assessment Database, and NHS Economic Evaluation Database

Database segments: EBM Reviews - Cochrane Central Register of Controlled Trials <August 2020>, EBM Reviews - Cochrane Database of Systematic Reviews <2005 to September 3, 2020>, EBM Reviews - Health Technology Assessment <4th Quarter 2016>, EBM Reviews - NHS Economic Evaluation Database <1st Quarter 2016>, Embase <1980 to 2020 Week 36>, Ovid MEDLINE(R) ALL <1946 to September 08, 2020>

Search strategy:

---

1 Thyroid Nodule/ (20379)
2 ((thyroid* or follicular*) adj3 (nodule* or mass* or cyst* or lesion*)).ti,ab,kf. (38070)
3 Thyroid Neoplasms/ (57825)
4 ((thyroid* or follicular*) adj2 (cancer* or neoplasm* or carcinoma* or adenoma* or adenocarcinoma* or tumo?r* or malignan* or metasta*)).ti,ab,kf. (117692)
5 or/1-4 (153112)
6 (atypi* or indetermin* or undetermin* or suspicio* or bethesda* or AUS or FLUS or FNSFN* or FN SFN* or TBSRTC* or BSRTC* or cytolog* or cytopatholog*).ti,ab,kf. (806053)
7 exp Biopsy, Fine-Needle/ (47593)
8 (fine needle* or aspirat* or biops* or FNA or FNAB or FNAC).ti,ab,kf. (1264977)
9 or/6-8 (1918604)
10 5 and 9 (33251)
11 Gene Expression Profiling/ (239334)
12 Gene Expression/ (1088212)
13 (afirma* or veracyte* or GEC or AGEC or GSC).ti,ab,kf. (8011)
14 classifier*.ti,ab,kf. (58080)
15 ((gene expression* or genomic expression*) adj2 (profil* or class* or pattern*)).ti,ab,kf. (138118)
16 ((genom* or genetic or gene) adj2 sequenc*).ti,ab,kf. (318862)
17 (thyroseq* or tqv2* or tqv3* or thygenx* or thygennext* or thyramir*).ti,ab,kf. (286)
18 ((miRNA* or microRNA* or mRNA*) adj2 (class* or express* or profil* or test*)).ti,ab,kf. (425146)
19 (gene mutat* panel* or GMP or ((multigene or multi gene) adj2 (panel* or test*))).ti,ab,kf. (39208)

20 (rosetta* or Dx15* or diaxonhit* or evicore*).ti,ab,kf. (4384)
21 Molecular Diagnostic Techniques/ (26875)
22 (molecular adj2 (class* or express* or profil* or test* or marker* or analy* or pattern* or cytolog* or diagnos* or trait*)).ti,ab,kf. (325067)
23 or/11-22 (2314988)
24 10 and 23 (3184)
25 exp Animals/ not Humans/ (16559315)
26 24 not 25 (2166)
27 Case Reports/ or Comment.pt. or Editorial.pt. or (Letter not (Letter and Randomized Controlled Trial)).pt. or Congress.pt. (5607098)
28 26 not 27 (2068)
29 limit 28 to yr="2019 -Current" (430)
30 limit 29 to english language [Limit not valid in CDSR; records were retained] (419)
31 30 use medall,cctr,coch,clhta,cleed (198)
32 thyroid nodule/ (20379)
33 ((thyroid* or follicular*) adj3 (nodule* or mass* or cyst* or lesion*)).tw,kw. (38405)
34 thyroid tumor/ (12913)
35 exp thyroid cancer/ (115964)
36 ((thyroid* or follicular*) adj2 (cancer* or neoplasm* or carcinoma* or adenoma* or adenocarcinoma* or tumo?r* or malignan* or metasta*)).tw,kw. (118007)
37 or/32-36 (168345)
38 (atypi* or indetermin* or undetermin* or suspicio* or bethesda* or AUS or FLUS or FNSFN* or FN SFN* or TBSRTC* or BSRTC* or cytolog* or cytopatholog*).tw,kw. (811421)
39 fine needle aspiration biopsy/ (40566)
40 (fine needle* or aspirat* or biops* or FNA or FNAB or FNAC).tw,kw. (1273705)
41 or/38-40 (1929394)
42 37 and 41 (34697)
43 gene expression profiling/ (239334)
44 gene expression/ (1088212)
45 (afirma* or veracyte* or GEC or AGEC or GSC).tw,kw,dv. (8081)
46 classifier*.tw,kw,dv. (58399)
47 ((gene expression* or genomic expression*) adj2 (profil* or class* or pattern*)).tw,kw,dv. (139545)
48 ((genom* or genetic or gene) adj2 sequenc*).tw,kw,dv. (320570)
49 (thyroseq* or tqv2* or tqv3* or thygenx* or thygennext* or thyramir*).tw,kw,dv. (298)
50 ((miRNA* or microRNA* or mRNA*) adj2 (class* or express* or profil* or test*)).tw,kw,dv. (426125)
51 (gene mutat* panel* or GMP or ((multigene or multi gene) adj2 (panel* or test*))).tw,kw,dv. (40170)
52 (rosetta* or Dx15* or diaxonhit* or evicore*).tw,kw,dv. (4507)
53 molecular diagnosis/ (19519)
54 (molecular adj2 (class* or express* or profil* or test* or marker* or analy* or pattern* or cytolog* or diagnos* or trait*)).tw,kw,dv. (328211)
55 or/43-54 (2314867)

56 42 and 55 (3261)
57 (exp animal/ or nonhuman/) not exp human/ (10769479)
58 56 not 57 (3240)
59 Case Report/ or Comment/ or Editorial/ or (letter.pt. not (letter.pt. and randomized controlled trial/)) or conference abstract.pt. or conference review.pt. (11396543)
60 58 not 59 (2041)
61 limit 60 to yr="2019 -Current" (435)
62 limit 61 to english language [Limit not valid in CDSR; records were retained] (421)
63 62 use emez (232)
64 31 or 63 (430)
65 64 use medall (197)
66 64 use emez (232)
67 64 use cctr (1)
68 64 use coch (0)
69 64 use clhta (0)
70 64 use cleed (0)
71 remove duplicates from 64 (255)

Economic Evidence Search

Search date: September 10, 2020

Databases searched: Ovid MEDLINE, Embase, Cochrane Central Register of Controlled Trials, Cochrane Database of Systematic Reviews, Centre for Reviews and Dissemination (CRD) Health Technology Assessment Database, and National Health Service (NHS) Economic Evaluation Database

Database segments: EBM Reviews - Cochrane Central Register of Controlled Trials <August 2020>, EBM Reviews - Cochrane Database of Systematic Reviews <2005 to September 3, 2020>, EBM Reviews - Health Technology Assessment <4th Quarter 2016>, EBM Reviews - NHS Economic Evaluation Database <1st Quarter 2016>, Embase <1980 to 2020 Week 36>, Ovid MEDLINE(R) ALL <1946 to September 09, 2020>

Search strategy:

---

1 Thyroid Nodule/ (20379)
2 ((thyroid* or follicular*) adj3 (nodule* or mass* or cyst* or lesion*)).ti,ab,kf. (38068)
3 Thyroid Neoplasms/ (57842)
4 ((thyroid* or follicular*) adj2 (cancer* or neoplasm* or carcinoma* or adenoma* or adenocarcinoma* or tumo?r* or malignan* or metasta*)).ti,ab,kf. (117682)
5 or/1-4 (153101)
6 (atypi* or indetermin* or undetermin* or suspicio* or bethesda* or AUS or FLUS or FNSFN* or FN SFN* or TBSRTC* or BSRTC* or cytolog* or cytopatholog*).ti,ab,kf. (805987)
7 exp Biopsy, Fine-Needle/ (47596)

8 (fine needle* or aspirat* or biops* or FNA or FNAB or FNAC).ti,ab,kf. (1264893)
9 or/6-8 (1918465)
10 5 and 9 (33250)
11 Gene Expression Profiling/ (239379)
12 Gene Expression/ (1088255)
13 (afirma* or veracyte* or GEC or AGEC or GSC).ti,ab,kf. (8008)
14 classifier*.ti,ab,kf. (58059)
15 ((gene expression* or genomic expression*) adj2 (profil* or class* or pattern*)).ti,ab,kf. (138096)
16 ((genom* or genetic or gene) adj2 sequenc*).ti,ab,kf. (318816)
17 (thyroseq* or tqv2* or tqv3* or thygenx* or thygennext* or thyramir*).ti,ab,kf. (286)
18 ((miRNA* or microRNA* or mRNA*) adj2 (class* or express* or profil* or test*)).ti,ab,kf. (425098)
19 (gene mutat* panel* or GMP or ((multigene or multi gene) adj2 (panel* or test*))).ti,ab,kf. (39206)
20 (rosetta* or Dx15* or diaxonhit* or evicore*).ti,ab,kf. (4382)
21 Molecular Diagnostic Techniques/ (26881)
22 (molecular adj2 (class* or express* or profil* or test* or marker* or analy* or pattern* or cytolog* or diagnos* or trait*)).ti,ab,kf. (325026)
23 or/11-22 (2314886)
24 10 and 23 (3184)
25 exp Animals/ not Humans/ (16559994)
26 24 not 25 (2166)
27 Case Reports/ or Comment.pt. or Editorial.pt. or (Letter not (Letter and Randomized Controlled Trial)).pt. or Congress.pt. (5606562)
28 26 not 27 (2068)
29 28 use coch,clhta,cleed (6)
30 economics/ (258370)
31 economics, medical/ or economics, pharmaceutical/ or exp economics, hospital/ or economics, nursing/ or economics, dental/ (876301)
32 economics.fs. (438213)
33 (econom* or price or prices or pricing or priced or discount* or expenditure* or budget* or pharmacoeconomic* or pharmaco-economic*).ti,ab,kf. (976628)
34 exp "costs and cost analysis"/ (607218)
35 (cost or costs or costing or costly).ti. (283124)
36 cost effective*.ti,ab,kf. (359130)
37 (cost* adj2 (util* or efficacy* or benefit* or minimi* or analy* or saving* or estimate* or allocation or control or sharing or instrument* or technolog*)).ab,kf. (234963)
38 models, economic/ (13883)
39 markov chains/ or monte carlo method/ (88162)
40 (decision adj1 (tree* or analy* or model*)).ti,ab,kf. (47378)
41 (markov or markow or monte carlo).ti,ab,kf. (141567)
42 quality-adjusted life years/ (43881)

43 (QOLY or QOLYs or HRQOL or HRQOLs or QALY or QALYs or QALE or QALEs).ti,ab,kf. (83533)
44 ((adjusted adj1 (quality or life)) or (willing* adj2 pay) or sensitivity analys*s).ti,ab,kf. (136970)
45 or/30-44 (2743716)
46 28 and 45 (158)
47 46 use medall,cctr (78)
48 or/29,47 (84)
49 limit 48 to english language [Limit not valid in CDSR; records were retained] (82)
50 thyroid nodule/ (20379)
51 ((thyroid* or follicular*) adj3 (nodule* or mass* or cyst* or lesion*)).tw,kw. (38403)
52 thyroid tumor/ (12913)
53 exp thyroid cancer/ (115981)
54 ((thyroid* or follicular*) adj2 (cancer* or neoplasm* or carcinoma* or adenoma* or adenocarcinoma* or tumo?r* or malignan* or metasta*)).tw,kw. (117996)
55 or/50-54 (168333)
56 (atypi* or indetermin* or undetermin* or suspicio* or bethesda* or AUS or FLUS or FNSFN* or FN SFN* or TBSRTC* or BSRTC* or cytolog* or cytopatholog*).tw,kw. (811354)
57 fine needle aspiration biopsy/ (40569)
58 (fine needle* or aspirat* or biops* or FNA or FNAB or FNAC).tw,kw. (1273623)
59 or/56-58 (1929256)
60 55 and 59 (34696)
61 gene expression profiling/ (239379)
62 gene expression/ (1088255)
63 (afirma* or veracyte* or GEC or AGEC or GSC).tw,kw,dv. (8078)
64 classifier*.tw,kw,dv. (58378)
65 ((gene expression* or genomic expression*) adj2 (profil* or class* or pattern*)).tw,kw,dv. (139524)
66 ((genom* or genetic or gene) adj2 sequenc*).tw,kw,dv. (320526)
67 (thyroseq* or tqv2* or tqv3* or thygenx* or thygennext* or thyramir*).tw,kw,dv. (298)
68 ((miRNA* or microRNA* or mRNA*) adj2 (class* or express* or profil* or test*)).tw,kw,dv. (426078)
69 (gene mutat* panel* or GMP or ((multigene or multi gene) adj2 (panel* or test*))).tw,kw,dv. (40168)
70 (rosetta* or Dx15* or diaxonhit* or evicore*).tw,kw,dv. (4505)
71 molecular diagnosis/ (19519)
72 (molecular adj2 (class* or express* or profil* or test* or marker* or analy* or pattern* or cytolog* or diagnos* or trait*)).tw,kw,dv. (328174)
73 or/61-72 (2314767)
74 60 and 73 (3261)
75 (exp animal/ or nonhuman/) not exp human/ (10770158)
76 74 not 75 (3240)
77 Case Report/ or Comment/ or Editorial/ or (letter.pt. not (letter.pt. and randomized controlled trial/)) or conference abstract.pt. or conference review.pt. (11396006)

78 76 not 77 (2041)
79 Economics/ (258370)
80 Health Economics/ or Pharmacoeconomics/ or Drug Cost/ or Drug Formulary/ (133545)
81 Economic Aspect/ or exp Economic Evaluation/ (476494)
82 (econom* or price or prices or pricing or priced or discount* or expenditure* or budget* or pharmacoeconomic* or pharmaco-economic*).tw,kw. (1003341)
83 exp "Cost"/ (607218)
84 (cost or costs or costing or costly).ti. (283124)
85 cost effective*.tw,kw. (371854)
86 (cost* adj2 (util* or efficac* or benefit* or minimi* or analy* or saving* or estimate* or allocation or control or sharing or instrument* or technolog*)).ab,kw. (247168)
87 Monte Carlo Method/ (69649)
88 (decision adj1 (tree* or analy* or model*)).tw,kw. (51247)
89 (markov or markow or monte carlo).tw,kw. (146643)
90 Quality-Adjusted Life Years/ (43881)
91 (QOLY or QOLYs or HRQOL or HRQOLs or QALY or QALYs or QALE or QALEs).tw,kw. (87442)
92 ((adjusted adj1 (quality or life)) or (willing* adj2 pay) or sensitivity analys*s).tw,kw. (158172)
93 or/79-92 (2359666)
94 78 and 93 (176)
95 94 use emez (99)
96 limit 95 to english language [Limit not valid in CDSR; records were retained] (97)
97 49 or 96 (179)
98 97 use medall (76)
99 97 use emez (97)
100 97 use cctr (0)
101 97 use coch (0)
102 97 use cleed (4)
103 97 use clhta (2)
104 remove duplicates from 97 (108)

Grey Literature Search

Performed: September 23–29, 2020

Websites searched: Alberta Health Evidence Reviews, Alberta Health Services, BC Health Technology Assessments, Canadian Agency for Drugs and Technologies in Health (CADTH), Institut national d'excellence en santé et en services sociaux (INESSS), Institute of Health Economics (IHE), McGill University Health Centre Health Technology Assessment Unit, Centre Hospitalier de l'Universite de Quebec-Universite Laval, Health Technology Assessment Database, Agency for Healthcare Research and Quality (AHRQ) Evidence-based Practice Centers, Centers for Medicare & Medicaid Services Technology Assessments, Veterans Affairs Health Services Research and Development, Institute for Clinical and Economic Review, Oregon

Health Authority Health Evidence Review Commission, Washington State Health Care Authority Health Technology Reviews, National Institute for Health and Care Excellence (NICE), Healthcare Improvement Scotland, Health Technology Wales, Ireland Health Information and Quality Authority Health Technology Assessments, Australian Government Medical Services Advisory Committee, Council of Australian Governments Health Technologies, Australian Safety and Efficacy Register of New Interventional Procedures -Surgical (ASERNIP-S), Italian National Agency for Regional Health Services (AGENAS), Belgian Health Care Knowledge Centre, Ludwig Boltzmann Institute for Health Technology Assessment, Swedish Agency for Health Technology Assessment and Assessment of Social Services, Ministry of Health Malaysia Health Technology Assessment Section, Tuft's Cost-Effectiveness Analysis Registry, PROSPERO, EUnetHTA, clinicaltrials.gov

Keywords used:
thyroid, thyroid cancer, thyroid nodule, thyroid lesion, follicular, fine needle aspiration, FNA, Bethesda, undetermined, indeterminate, atypia, Afirma, Veracyte, gene expression classifier, gene expression, classifier, genetic testing, ThyroSeq, ThyGenX, Thygennext, Thyramir, molecular profiling

Clinical results (included in PRISMA): 1
Economic results (included in PRISMA): 6
Ongoing HTAs (PROSPERO/EUnetHTA/MSAC): 4
Ongoing RCTs (clinicaltrials.gov): 12

Appendix 2: Critical Appraisal of Clinical Evidence

Table A1: Risk of Bias[a] Among Systematic Reviews (ROBIS Tool)

	Identifying concerns with the review process				Judging risk of bias
Author, year	Study eligibility criteria	Identification and selection of studies	Data collection and study appraisal	Synthesis and findings	Risk of bias in the review
Borowczyk et al, 2019[128,129]	Low	Low	Low	Low	Low
INESSS, 2021[18]	Low	Low	Low	Low	Low
Liu et al, 2019[130]	Low	High	High	High	High
Muzza et al, 2020[131]	Low	High	High	High	High
Ngo et al, 2020[132]	Low	Low	High[b]	Low	Low
Valderrabano et al, 2019[133]	Low	Low	Low	Low	Low
Vuong et al, 2020[134]	Low	Low	Low	Low	Low

Abbreviation: INESSS, Institut national d'excellence en santé et en services sociaux; ROBIS, Risk of Bias in Systematic Reviews.

[a] Possible risk-of-bias levels: low, high, unclear.

[b] Rated high because of a lack of double reviewer assessment; however, given the totality of the methodology applied, this issue was not deemed a substantial concern, so the review was assessed as having a low risk of bias overall.

Table A2: Risk of Bias[a] Among Diagnostic Accuracy Studies (QUADAS-2 Tool)—Clinical Validity[b]

	Risk of bias				Applicability concerns		
Author, year	Patient selection	Index test	Reference standard	Flow and timing	Patient selection	Index test	Reference standard
Patel et al, 2018[54]	High[c]	Low	Unclear[d]	Low[e]	High[c]	Low	Low
Steward et al, 2019[56]	High[c]	Low	Unclear[d]	Low[e]	High[c]	Low	Low

Abbreviation: INESSS, Institut national d'excellence en santé et en services sociaux; QUADAS, Quality Assessment of Diagnostic Accuracy Studies.

[a] Possible risk-of-bias levels: low, high, unclear.

[b] Table completed using the INESSS report[18] as a source, supplemented by referral to the primary studies as needed.

[c] Eligibility was generally limited to people who were available for surgical follow-up, and may have introduced bias. The study population was judged to not necessarily be representative of the target population for the intervention (those eligible for first-line nodule assessment).

[d] Follow-up for people who did not undergo surgery was unspecified and unclear.

[e] The number of nodules excluded from the analyses was unclear; however, the overall risk of bias for flow and timing was considered low.

Table A3: Risk of Bias[a] Among Non-randomized Studies (RoBANS Tool)[b]—Clinical Utility

Author, year	Selection of participants	Confounding variables	Measurement of exposure	Blinding of outcome assessments	Incomplete outcome data
Angell et al, 2019[51]	High[c,d]	Unclear[e]	Low	Low	Low
Chen et al, 2020[33]	High[c,d]	Unclear[e]	Low	Low	Low
Endo et al, 2019[52]	High[c]	Unclear[e]	Low	Low	High[f]
Harrell et al, 2019[53]	High[c]	Unclear[e]	Low	Low	Low
San Martin et al, 2020[55]	High[c]	Unclear[e]	Low	Low	Low
Wei et al, 2019[57]	High[c]	Unclear[e]	Low	Low	Low

Abbreviation: INESSS, Institut national d'excellence en santé et en services sociaux; RoBANS, Risk of Bias Assessment Tool for Non-randomized Studies.

[a] Possible risk-of-bias levels: low, high, unclear.

[b] Table completed using the INESSS report[18] as a source, supplemented by referral to the primary studies as needed.

[c] It was unclear whether the people who received molecular testing were the same as those who did not; studies were not randomized, and there was no information about the group that did not receive molecular testing.

[d] Whether to receive molecular testing, active surveillance, or surgery was a choice for patients to make after discussion with their clinician.

[e] The study design did not allow for adequate comparison to a group that did not receive molecular testing.

[f] The proportion of potentially eligible nodules that were excluded for a variety of reasons was high enough that this missing data could have affected study outcomes.

Table A4: GRADE Evidence Profile for the Evaluation of Molecular Testing of Thyroid Nodules of Indeterminate Cytology[a]

Number of studies (design)	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Upgrade considerations	Quality
Clinical validity (diagnostic accuracy)							
Afirma GSC 1 (diagnostic accuracy)	Serious limitations (−1)[b]	No serious limitations	No serious limitations[c]	No serious limitations	Likely (−1)[d]	None	⨁⨁ Low
ThyroSeq v3 1 (diagnostic accuracy)	Serious limitations (−1)[b]	No serious limitations	No serious limitations[c]	No serious limitations	Likely (−1)[d]	None	⨁⨁ Low
Clinical utility (resection rate)							
Afirma GSC 5 (observational)	Serious limitations (−1)[f]	No serious limitations	No serious limitations[c]	Not calculable[g]	Likely (−1)[d]	None	⨁ Very Low
ThyroSeq v3 1 (observational)	Serious limitations (−1)[f]	No serious limitations	No serious limitations[c]	Not calculable[g]	Strongly suspected (−1)[d,e]	None	⨁ Very Low

Abbreviations: GRADE, Grading of Recommendations Assessment, Development, and Evaluation; GSC, gene sequencing classifier; INESSS, Institut national d'excellence en santé et en services sociaux.

[a] Table completed using the INESSS report[18] as a source, supplemented by referral to the primary studies as needed.

[b] Details of risk-of-bias assessment in Table A2.

[c] The studies were designed in a way that could be reasonably assumed to reflect real-world clinical practice.

[d] Publication bias was a concern because of the low number of publications, and because most had industry sponsorship, either directly or as a conflict of interest declared by authors.

[e] Only one publication had positive findings on a new technology, leading to a concern of possible lag bias.[135]

[f] Details of risk-of-bias assessment in Table A3.

[e] Resection rates were reported as yields, and precision was not calculable.

Appendix 3: Study Characteristics—Clinical Evidence

Table A5: Molecular Testing of Thyroid Nodules of Indeterminate Cytology—Characteristics of Systematic Reviews

Author, year Search dates	Population	Intervention	Studies, N
Borowczyk et al, 2019[128,129] January 2001 to April 2018	Indeterminate nodules (TBSRTC III or IV); excluded cases without postsurgical diagnosis and those that assessed only AUS without FLUS	GEC or ThyroSeq	16
INESSS, 2021[18] Literature search on March 19, 2019, followed by a periodic search for updated evidence until publication date	Indeterminate nodules (TBSRTC III or IV)	Any commercially available rule-out molecular test	46
Liu et al, 2019[130] January 2005 to December 2018	Indeterminate nodules (including TBSRTC III, IV, or V); excluded incidental microcarcinomas	GEC	18
Muzza et al, 2020[131] 2009 to 2019	Indeterminate nodules (TBSRTC III or IV); studies with more than 20 cases	Rule-in and rule-out panels with more than four genes or miRNA; included both commercially available and non-commercial panels	32
Ngo et al, 2020[132] January 2010 to March 2019	Indeterminate nodules (TBSRTC III or IV)	Molecular tests performed preoperatively for nodules meeting the NIFTP criteria requiring surgery or equal to malignancy	33
Valderrabano et al, 2019[133] Up to October 26, 2017	Nodules with AUS/FLUS or FN cytology results with sufficient RNA for GEC testing; studies with at least 35 nodules assessed	GEC	19
Vuong et al, 2020[134] Up to April 30, 2020	Indeterminate nodules (not specified)	GEC and GSC	7

Abbreviation: AUS, atypical lesion of undetermined significance; FLUS, follicular lesion of undetermined significance; FN, follicular neoplasm; GEC, gene expression classifier; GSC, gene sequencing classifier; miRNA, microRNA; NIFTP, noninvasive follicular thyroid neoplasm; TBSRTC, The Bethesda System for Reporting Thyroid Cytopathology.

Appendix 4: Selected Excluded Studies—Economic Evidence

For transparency, we provide a list of studies that readers might have expected to see but that did not meet the inclusion criteria, along with the primary reasons for exclusion. These economic analyses yielded mixed results: some showed molecular testing to be cost-saving[86,89,90,94] or cost-effective,[84] and others showed the opposite.[83,87,88,91,92] Factors that may have contributed to the variation in results include analysis approach, test cost, surgery cost, and the prevalence of malignancy.

Table A6: List of Excluded Economic Studies and Reasons for Exclusion

Author, year, country	Reasons for Exclusion
Zanocco et al, 2020,[87] United States	• Evaluated an earlier version of the molecular tests (e.g., Afirma GEC) • The target population was stratified into 3 groups: high, intermediate, and low sonographic suspicion
Fazeli et al, 2020,[88] United States	• Evaluated an earlier version of the molecular tests (e.g., ThyroSeq v2) • Cost analysis only (based on a cohort study before and after the introduction of molecular testing)
Balentine et al, 2018,[83] United States	• Evaluated an earlier version of the molecular tests (e.g., Afirma GEC) • Evaluated the "rule-in" use of molecular testing (e.g., some patients with an Afirma GEC "suspicious" result would get a total thyroidectomy, and some would get a diagnostic lobectomy) • Usual care in the reference case analysis was based on older guidelines (the 2009 American Thyroid Association guidelines[81] recommended a completion thyroidectomy for all cancers ≥ 1 cm detected after diagnostic lobectomy). The model evaluated the impact of following the 2015 American Thyroid Association guidelines[9] in a scenario analysis: lobectomy constitutes adequate treatment for most thyroid cancers in the size range of 1 to 4 cm
Rivas et al, 2018,[89] United States	• Evaluated an earlier version of the molecular tests (e.g., ThyroSeq v2) • Cost analysis only (based on a cohort study)
Shapiro et al, 2017,[91] United States	• Evaluated an earlier version of the molecular tests (e.g., Afirma GEC) • Model was based on the 2009 American Thyroid Association guidelines[81]
Abeykoon et al, 2016,[90] United States	• Evaluated an earlier version of the molecular tests (e.g., Afirma GEC) • Cost analysis only (based on a cohort study)
Labourier, 2016,[84] United States	• Evaluated an earlier version of the molecular tests (e.g., Afirma GEC) • Evaluated the "rule-in" use of molecular testing (test result used to determine the extent of surgery from diagnostic lobectomy to total thyroidectomy) • Model assumed that all malignant nodules detected after diagnostic lobectomy would receive a completion thyroidectomy. However, current guidelines recommend completion thyroidectomy only in specific cases
Singer et al, 2016,[136] United States	• Did not compare molecular testing to usual care; instead compared cytopathology benign patients to Afirma GEC benign patients

Author, year, country	Reasons for Exclusion
Wu et al, 2016,[92] United States	• Evaluated an earlier version of the molecular tests (e.g., Afirma GEC) • Model was based on the 2009 American Thyroid Association guidelines[81]
Lee et al, 2014,[93] Canada and United States	• Evaluated an earlier version of the molecular tests (e.g., Afirma GEC) • Model was based on the 2009 American Thyroid Association guidelines[81]
Vriens et al, 2014,[85] Netherlands	• Evaluated the "rule-in" use of molecular testing (e.g., ThyroSeq v2 [MMP] was used to determine the extent of surgery) • Evaluated earlier versions of the molecular tests (e.g., Afirma GEC; ThyroSeq v2)
Najafzadeh et al, 2012,[137] United States	• Evaluated a hypothetical molecular test • Included thyroid nodules with TBSRTC cytology I to VI • Molecular testing was not compared to diagnostic lobectomy, but to a comprehensive approach that combined biopsy with classification according to TBSRTC
Yip et al, 2012,[86] United States	• Examined a preliminary version of ThyroSeq used as rule-in test • Included thyroid nodules with TBSRTC cytology II and V • Cost analysis only
Li et al, 2011,[94] United States	• Evaluated a preliminary version of the Afirma GEC (Chudova et al, 2010[138]) • Model was based on the 2009 American Thyroid Association guidelines[81]

Abbreviations: GEC, gene expression classifier; MMP, xx; TBSRTC, The Bethesda System for Reporting Thyroid Cytology.

Appendix 5: Results of Applicability Checklist for Studies Included in the Economic Literature Review

Table A7: Assessment of the Applicability of Studies Evaluating the Cost-Effectiveness of Molecular Testing for Thyroid Nodules of Indeterminate Cytology

Author, Year, Country	Is the study population similar to the question?	Are the interventions similar to the question?	Is the health care system studied sufficiently similar to Ontario?	Were the perspectives clearly stated? If yes, what were they?	Are all direct effects included? Are all other effects included where they are material?	Are all future costs and outcomes discounted? If yes, at what rate?	Is the value of health effects expressed in terms of quality-adjusted life-years?	Are costs and outcomes from other sectors fully and appropriately measured and valued?	Overall judgment[a]
Nicholson et al, 2019,[95] United States	Yes	Yes	No	Yes, US payer perspective	No (did not consider effects on patient's quality of life or different resection rates following a positive or negative test result)	Partially (only cost discounted at 3% per year)	No (expressed as the probability of predicting a correct diagnosis)	No (did not consider costs borne by patients and society)	Not applicable
INESSS, 2021,[18] Quebec, Canada	Yes	Yes	Partially (some key resource use and cost parameters were different)	Yes, Québec public payer perspective	No (did not consider test failure rate or effect on the rate of surgery)	Not reported	Yes	No (did not consider costs borne by patients and society)	Partially applicable

Abbreviation: INESSS, Institut national d'excellence en santé et en services sociaux.

Note: Response options for all items were "yes," "partially," "no," "unclear," and "NA" (not applicable).

[a] Overall judgment may be "directly applicable," "partially applicable," or "not applicable."

Appendix 6: Cost-Effectiveness Plane Scatterplots

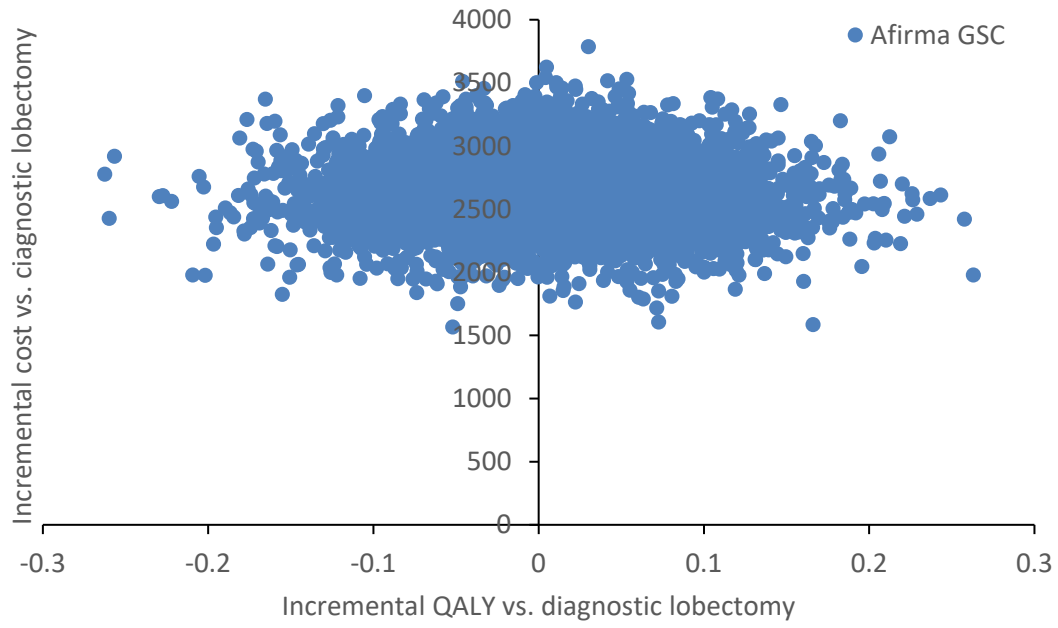

Figure A1: Cost-Effectiveness Scatterplot—Afirma GSC Versus Diagnostic Lobectomy

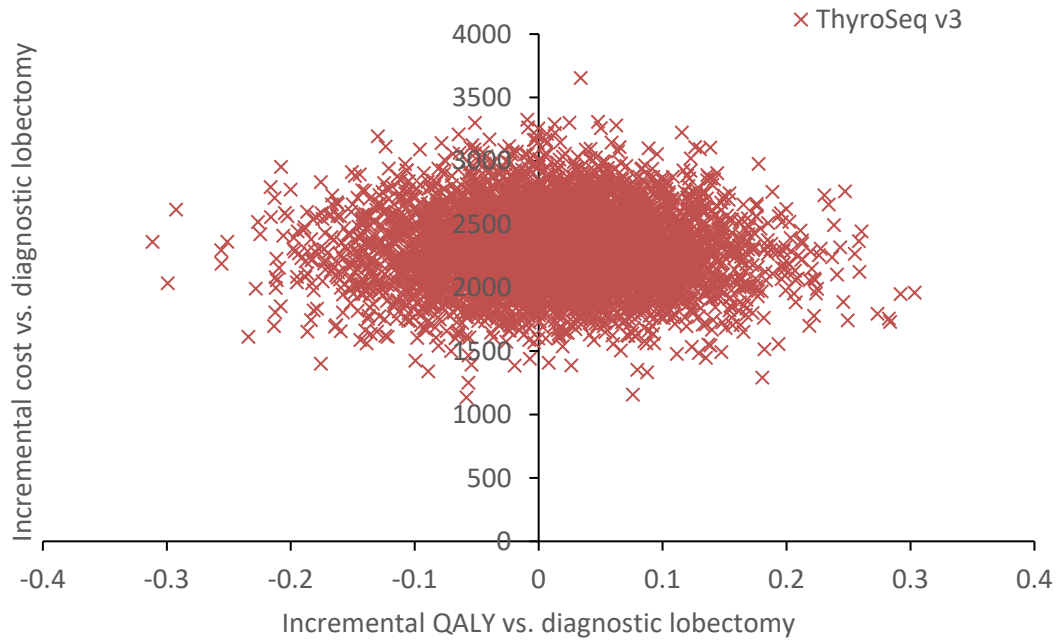

Figure A2: Cost-Effectiveness Scatterplot—ThyroSeq v3 Versus Diagnostic Lobectomy

Appendix 7: Letter of Information

Ontario Health

LETTER OF INFORMATION

Ontario Health is conducting a review of **molecular testing of thyroid nodules**. The purpose is to understand whether this testing should be more broadly funded in Ontario.

An important part of this review involves speaking to patients and family members of those who may have experience with thyroid nodules and molecular testing, or who may have attempted to access it. Our goal is always to make sure the lived-experience of individuals and families are considered in the funding recommendations for this test.

WHAT DO YOU NEED FROM PARTICIPANTS?

- ✓ 20-30 minutes of time for a phone or in-person interview to hear about their experiences
- ✓ Permission to audio- (not video-) record the interview

WHAT PARTICIPATION INVOLVES

If a participant agrees to share their experiences, they will be asked to have an interview with Ontario Health staff. The interview will likely last 20-30 minutes. It will be held in a private location or over the telephone. With consent, the interview will be audio-recorded. The interviewer will ask questions about perspectives of thyroid nodules, decision-making and more general thoughts about molecular testing and thyroid treatment options in Ontario.

Participation is voluntary. Those who volunteer may decide not to participate, refuse to answer any questions or withdraw before the interview. Withdrawal will in no way affect the care received.

CONFIDENTIALITY

All information collected for the review will be kept confidential and privacy will be protected except as required by law. The results of this review will be published, however no identifying information will be released or published. Any records containing information from the interview will be stored securely.

RISKS TO PARTICIPATION:

There are no known physical risks to participating. Some participants may experience discomfort or anxiety after speaking about their lived experience. If this is the case, participants can speak to our staff.

If you have any questions, please contact Ontario Health staff:

Appendix 8: Interview Guide

Interview for Thyroid Genetic Testing HTA

Intro
Explain HQO purpose, HTA process, and purpose of interview

Lived- Experience
Any symptoms that led to discovery of thyroid nodules?
Impact of symptoms on quality of life?
Medical journey to receiving diagnosis
Access to diagnosis and specialists – any barriers?

Treatments and Decision-Making
What kind of info was provided about thyroid nodules and potential for development of cancer?
Where did this info come from (doc, friends, research, etc)
Decision-making surrounding different treatment options – what factors were important to you?

Genetic Testing
Would information from genetic testing be valuable to you?
What (if anything) would make you hesitate about getting genetic testing?
Procedure? Cost?
Genetic testing and decision-making – would it influence it?

Overall thoughts on the use of genetic testing for thyroid nodules and its potential impact?

References

(1) Mummadi SR, de Longpre J, Hahn PY. Comparative effectiveness of interventions in initial management of spontaneous pneumothorax: a systematic review and a Bayesian network meta-analysis. Ann Emerg Med. 2020;76(1):88-102.

(2) Patell R, Mikhael A, Tabet M, Bena J, Berber E, Nasr C. Assessing the utility of preoperative serum thyroglobulin in differentiated thyroid cancer: a retrospective cohort study. Endocrine. 2018;61(3):506-10.

(3) Shah BR, Griffiths R, Hall SF. Thyroid cancer incidence among Asian immigrants to Ontario, Canada: a population-based cohort study. Cancer. 2017;123(17):3320-5.

(4) Topstad D, Dickinson JA. Thyroid cancer incidence in Canada: a national cancer registry analysis. CMAJ Open. 2017;5(3):E612-E6.

(5) Ally M, Patel N, Bhatia K, Awad Z. Papillary thyroid cancer nodal metastases mimicking a slow-flow lymphatic malformation. BMJ Case Rep. 2018;2018.

(6) Paschke R, Cantara S, Crescenzi A, Jarzab B, Musholt TJ, Sobrinho Simoes M. European Thyroid Association Guidelines regarding Thyroid Nodule Molecular Fine-Needle Aspiration Cytology Diagnostics. Eur Thyroid J. 2017;6(3):115-29.

(7) Eskander A, Monteiro E, O'Connell D, Taylor SM, Canadian Association of Head and Neck Surgical Oncology. Head and Neck Surgical Oncology Choosing Wisely Campaign: imaging for patients with hoarseness, fine needle aspiration for neck mass, and ultrasound for odynophagia. J Otolaryngol Head Neck Surg. 2018;47(1):2.

(8) Davies L, Welch HG. Current thyroid cancer trends in the United States. JAMA Otolaryngol Head Neck Surg. 2014;140(4):317-22.

(9) Haugen BR, Alexander EK, Bible KC, Doherty GM, Mandel SJ, Nikiforov YE, 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.

(10) Moubayed SP, Urken ML. Thyroid nodules. CMAJ. 2016;188(17-18):1259.

(11) Cancer Care Ontario. Ontario cancer statistics 2018 [Internet]. Toronto: Cancer Care Ontario; 2018 [cited 2021 May 27]. Available from: https://www.cancercareontario.ca/sites/ccocancercare/files/assets/OCS2018_rev13122018.pdf

(12) Oczko-Wojciechowska M, Kotecka-Blicharz A, Krajewska J, Rusinek D, Barczynski M, Jarzab B, et al. European perspective on the use of molecular tests in the diagnosis and therapy of thyroid neoplasms. Gland Surg. 2020;9(Suppl 2):S69-S76.

(13) Verloop H, Louwerens M, Schoones JW, Kievit J, Smit JW, Dekkers OM. Risk of hypothyroidism following hemithyroidectomy: systematic review and meta-analysis of prognostic studies. J Clin Endocrinol Metab. 2012;97(7):2243-55.

(14) Yoo J, Agbassi C, Lochnan H, Morrison D, Sawka A, Brierley J, et al. Cancer Care Ontario Thyroid Cancer Guideline: an endorsement of the 2015 American Thyroid Association management guidelines for adult patients with thyroid nodules and differentiated thyroid cancer. Program in evidence-based care guideline no. 5-13. [Internet]. Toronto: Cancer Care Ontario; 2017 [cited 2021 Jun 29]. Available from: https://www.partnershipagainstcancer.ca/db-sage/sage20180092/

(15) Haugen BR. 2015 American Thyroid Association management guidelines for adult patients with thyroid nodules and differentiated thyroid cancer: what is new and what has changed? Cancer. 2017;123(3):372-81.

(16) Bongiovanni M, Spitale A, Faquin WC, Mazzucchelli L, Baloch ZW. The Bethesda system for reporting thyroid cytopathology: a meta-analysis. Acta Cytol. 2012;56(4):333-9.

(17) Cibas ES, Ali SZ. The 2017 Bethesda system for reporting thyroid cytopathology. Thyroid. 2017;27(11):1341-6.
(18) Institut national d'excellence en santé et en services sociaux (INESSS). Pertinence de l'utilisation des tests moléculaires pour préciser le diagnostic des nodules thyroïdiens à cytologie indéterminée [Internet]. Québec (QC): INESSS; 2021 [cited 2021 Feb 28]. Available from: https://www.inesss.qc.ca/fileadmin/doc/INESSS/Rapports/Technologies/INESSS_ThyroSeq_Avis.pdf
(19) Manning AM, Yang H, Falciglia M, Mark JR, Steward DL. Thyroid ultrasound-guided fine-needle aspiration cytology results: observed increase in indeterminate rate over the past decade. Otolaryngol Head Neck Surg. 2017;156(4):611-5.
(20) Fernandes VT, De Santis RJ, Enepekides DJ, Higgins KM. Surgeon-performed ultrasound guided fine-needle aspirate biopsy with report of learning curve: a consecutive case-series study. J Otolaryngol Head Neck Surg. 2015;44:42.
(21) Fernandes VT, Magarey MJ, Kamdar DP, Freeman JL. Surgeon performed ultrasound-guided fine-needle aspirates of the thyroid: 1067 biopsies and learning curve in a teaching center. Head Neck. 2016;38 Suppl 1:E1281-4.
(22) Allen L, Al Afif A, Rigby MH, Bullock MJ, Trites J, Taylor SM, et al. The role of repeat fine needle aspiration in managing indeterminate thyroid nodules. J Otolaryngol Head Neck Surg. 2019;48(1):16.
(23) Bernstein JM, Shah M, MacMillan C, Freeman JL. Institution-specific risk of papillary thyroid carcinoma in atypia/follicular lesion of undetermined significance. Head Neck. 2016;38 Suppl 1:E1210-5.
(24) Baloch ZW, LiVolsi VA, Asa SL, Rosai J, Merino MJ, Randolph G, et al. Diagnostic terminology and morphologic criteria for cytologic diagnosis of thyroid lesions: a synopsis of the National Cancer Institute Thyroid Fine-Needle Aspiration State of the Science Conference. Diagn Cytopathol. 2008;36(6):425-37.
(25) Nishino M, Krane JF. Role of ancillary techniques in thyroid cytology specimens. Acta Cytologica. 2020;64(1-2):40-51.
(26) Krane JF, Cibas ES, Endo M, Marqusee E, Hu MI, Nasr CE, et al. The Afirma Xpression Atlas for thyroid nodules and thyroid cancer metastases: Insights to inform clinical decision-making from a fine-needle aspiration sample. Cancer Cytopathol. 2020;128(7):452-9.
(27) Rao SN, Bernet V. Indeterminate thyroid nodules in the era of molecular genomics. Mol Genet Genomic Med. 2020:e1288.
(28) Cabanillas ME, Dadu R, Iyer P, Wanland KB, Busaidy NL, Ying A, et al. Acquired secondary RAS mutation in BRAF(V600E)-mutated thyroid cancer patients treated with BRAF inhibitors. Thyroid. 2020;30(9):1288-96.
(29) Chu YH, Sadow PM. Noninvasive follicular thyroid neoplasm with papillary-like nuclear features (NIFTP): diagnostic updates and molecular advances. Semin Diagn Pathol. 2020;37(5):213-8.
(30) Pusztaszeri M, Bongiovanni M. The impact of non-invasive follicular thyroid neoplasm with papillary-like nuclear features (NIFTP) on the diagnosis of thyroid nodules. Gland Surg. 2019;8(Suppl 2):S86-S97.
(31) Rosario PW, Ward LS, Graf H, Vaisman F, Mourao GF, Vaisman M. Thyroid nodules </= 1 cm and papillary thyroid microcarcinomas: Brazilian experts opinion. Arch Endocrinol Metab. 2019;63(5):456-61.
(32) Kay-Rivest E, Tibbo J, Bouhabel S, Tamilia M, Leboeuf R, Forest VI, et al. The first Canadian experience with the Afirma(R) gene expression classifier test. J Otolaryngol Head Neck Surg. 2017;46(1):25.

(33) Chen T, Gilfix BM, Rivera J, Sadeghi N, Richardson K, Hier MP, et al. The role of the ThyroSeq v3 molecular test in the surgical management of thyroid nodules in the Canadian public health care setting. Thyroid. 2020;30(9):1280-7.
(34) Harrison G, Sosa JA, Jiang X. Evaluation of the Afirma gene expression classifier in repeat indeterminate thyroid nodules. Arch Pathol Lab Med. 2017;141(7):985-9.
(35) Roth MY, Witt RL, Steward DL. Molecular testing for thyroid nodules: review and current state. Cancer. 2018;124(5):888-98.
(36) Berger ML, Sox H, Willke RJ, Brixner DL, Eichler HG, Goettsch W, et al. Good practices for real-world data studies of treatment and/or comparative effectiveness: recommendations from the joint ISPOR-ISPE Special Task Force on real-world evidence in health care decision making. Pharmacoepidemiol Drug Saf. 2017;26(9):1033-9.
(37) Ontario Health (Cancer Care Ontario). Pandemic planning clinical guideline for patients with cancer [Internet]. Toronto: Ontario Health (Cancer Care Ontario); 2020 [cited 2020 Jun 29]. Available from: https://www.entcanada.org/wp-content/uploads/OH-CCO-Pandemic-Planning-Clinical-Guideline-Final.pdf
(38) Interpace diagnostics announces additional coverage of thyroid tests [Internet]. Parsippany (NJ): Interpace Biosciences; 2018 [updated 2020 Jun 16; cited 2020 Jun 29]. Available from: https://ir.interpace.com/press-releases/detail/86/interpace-diagnostics-announces-additional-coverage-of
(39) Afirma. In network insurance plans [Internet]. San Francisco (CA): Afirma; 2018 [updated 2018; cited 2020 Nov 19]. Available from: https://www.afirma.com/physicians/practice-resources/
(40) ThyroSeq. Insurance coverage and CPT code [Internet]. Atlanta (GA): ThyroSeq; 2017 [updated 2017; cited 2020 Nov 19]. Available from: https://thyroseq.com/physicians/test-ordering/insurance-coverage
(41) Chen CC, Hang JF, Liu CY, Wang YH, Lai CR. Thyroid fine-needle aspiration cytology in Taiwan: a nationwide survey and literature update. J Pathol Transl Med. 2020;54(5):361-6.
(42) Rosenblum RC, Shtabsky A, Marmor S, Trejo L, Yaish I, Barnes S, et al. Indeterminate nodules by the Bethesda system for reporting thyroid cytopathology in Israel: frequency, and risk of malignancy after reclassification of follicular thyroid neoplasm with papillary-like features. Eur J Surg Oncol. 2019;45(7):1182-7.
(43) Gomez-Perez AM, Fernandez-Garcia JC, Iglesias P, Diez JJ, Alvarez-Escola C, Lecumberri B, et al. Diagnosis and treatment of thyroid nodules in Spain. Results of a national survey. Endocrinol Diabetes Nutr. 2020;67(7):438-45.
(44) O'Neill J, Tabish H, Welch V, Petticrew M, Pottie K, Clarke M, et al. Applying an equity lens to interventions: using PROGRESS ensures consideration of socially stratifying factors to illuminate inequities in health. J Clin Epidemiol. 2014;67(1):56-64.
(45) Covidence [Computer program]. Melbourne (Australia): Veritas Health Innovation; 2021.
(46) Whiting P, Savovic J, Higgins JP, Caldwell DM, Reeves BC, Shea B, et al. ROBIS: a new tool to assess risk of bias in systematic reviews was developed. J Clin Epidemiol. 2016;69:225-34.
(47) Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med. 2011;155(8):529-36.
(48) Kim SY, Park JE, Lee YJ, Seo HJ, Sheen SS, Hahn S, et al. Testing a tool for assessing the risk of bias for nonrandomized studies showed moderate reliability and promising validity. J Clin Epidemiol. 2013;66(4):408-14.
(49) Marti JL, Avadhani V, Donatelli LA, Niyogi S, Wang B, Wong RJ, et al. Wide inter-institutional variation in performance of a molecular classifier for indeterminate thyroid nodules. Ann Surg Oncol. 2015;22(12):3996-4001.

(50) Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med. 2009;6(6):e1000097.
(51) Angell TE, Heller HT, Cibas ES, Barletta JA, Kim MI, Krane JF, et al. Independent comparison of the Afirma genomic sequencing classifier and gene expression classifier for cytologically indeterminate thyroid nodules. Thyroid. 2019;29(5):650-6.
(52) Endo M, Nabhan F, Porter K, Roll K, Shirley LA, Azaryan I, et al. Afirma gene sequencing classifier compared with gene expression classifier in indeterminate thyroid nodules. Thyroid. 2019;29(8):1115-24.
(53) Harrell RM, Eyerly-Webb SA, Golding AC, Edwards CM, Bimston DN. Statistical comparison of Afirma GSC and Afirma GEC outcomes in a community endocrine surgical practice: early findings. Endocrine Practice. 2019;25(2):161-4.
(54) Patel KN, Angell TE, Babiarz J, Barth NM, Blevins T, Duh QY, et al. Performance of a genomic sequencing classifier for the preoperative diagnosis of cytologically indeterminate thyroid nodules. JAMA Surg. 2018;153(9):817-24.
(55) San Martin VT, Lawrence L, Bena J, Madhun NZ, Berber E, Elsheikh TM, et al. Real-world Comparison of Afirma GEC and GSC for the assessment of cytologically indeterminate thyroid nodules. J Clin Endocrinol Metab. 2020;105(3):01.
(56) Steward DL, Carty SE, Sippel RS, Yang SP, Sosa JA, Sipos JA, et al. Performance of a multigene genomic classifier in thyroid nodules with indeterminate cytology: a prospective blinded multicenter study. JAMA Oncol. 2019;5(2):204-12.
(57) Wei S, Veloski C, Sharda P, Ehya H. Performance of the Afirma genomic sequencing classifier versus gene expression classifier: an institutional experience. Cancer Cytopathology. 2019;127(11):720-4.
(58) Duh QY, Busaidy NL, Rahilly-Tierney C, Gharib H, Randolph G. A systematic review of the methods of diagnostic accuracy studies of the Afirma gene expression classifier. Thyroid. 2017;27(10):1215-22.
(59) Schueler S, Schuetz GM, Dewey M. The revised QUADAS-2 tool. Ann Intern Med. 2012;156(4):323-4.
(60) National Library of Medicine. Molecular classifier for the fine needle-based assessment of malignancy risk in thyroid nodules (ThyroPred-1) [Internet]. Bethesda (MD): The Library; 2019 [updated 2019; cited 2021 Apr 7]. Available from: https://clinicaltrials.gov/ct2/show/NCT03392402?term=molecular+testing+thyroid&draw=2&rank=2
(61) Kant R, Davis A, Verma V. Thyroid nodules: advances in evaluation and management. Am Fam Physician. 2020;102(5):298-304.
(62) Li W, Justice-Clark T, Cohen MB. The utility of ThyroSeq((R)) in the management of indeterminate thyroid nodules by fine-needle aspiration. Cytopathology. 2021;32(4):505-12.
(63) Polavarapu P, Fingeret A, Yuil-Valdes A, Olson D, Patel A, Shivaswamy V, et al. Comparison of Afirma GEC and GSC to nodules without molecular testing in cytologically indeterminate thyroid nodules. J Endocr Soc. 2021;5(11):bvab148.
(64) Giorgenon TMV, Carrijo FT, Arruda MA, Cerqueira TLO, Barreto HR, Cabral JB, et al. Preoperative detection of TERT promoter and BRAFV600E mutations in papillary thyroid carcinoma in high-risk thyroid nodules. Arch Endocrinol Metab. 2019;63(2):107-12.
(65) Gucer H, Caliskan S, Kefeli M, Mete O. Do you know the details of your PAX8 antibody? Monoclonal PAX8 (MRQ-50) is not expressed in a series of 45 medullary thyroid carcinomas. Endocr Pathol. 2020;31(1):33-8.

(66) Mahmoudian-Sani MR, Amrollahi-Sharifabadi M, Taheri A, Hosseini SM, Tahmasebi K, Mobini GR. Diagnostic value of the candidate microRNAs in thyroid fine-needle aspiration biopsy (FNAB) samples. Horm Mol Biol Clin Investig. 2020;41(2).
(67) Samija I, Matesa N, Kozaj S, Majstorovic I, Bolanca A, Kusic Z. HMGA2 gene expression in fine-needle aspiration samples of thyroid nodules as a marker for preoperative diagnosis of thyroid cancer. Appl Immunohistochem Mol Morphol. 2019;27(6):471-6.
(68) Wu D, Hu S, Hou Y, He Y, Liu S. Identification of potential novel biomarkers to differentiate malignant thyroid nodules with cytological indeterminate. BMC Cancer. 2020;20(1):199.
(69) Xu X, Ma X, Zhang X, Cao G, Tang Y, Deng X, et al. Detection of BRAF V600E mutation in fine-needle aspiration fluid of papillary thyroid carcinoma by droplet digital PCR. Clin Chim Acta. 2019;491:91-6.
(70) Yim JH, Choi AH, Li AX, Qin H, Chang S, Tong ST, et al. Identification of tissue-specific DNA methylation signatures for thyroid nodule diagnostics. Clin Cancer Res. 2019;25(2):544-51.
(71) De Luca C, Sgariglia R, Nacchio M, Pisapia P, Migliatico I, Clery E, et al. Rapid on-site molecular evaluation in thyroid cytopathology: a same-day cytological and molecular diagnosis. Diagn Cytopathol. 2020;48(4):300-7.
(72) Kumar G, Timmaraju VA, Song-Yang JW, Repko B, Narick C, Mireskandari A, et al. Utility of microdissected cytology smears for molecular analysis of thyroid malignancy. Diagn Cytopathol. 2019;47(4):289-96.
(73) Khatami F, Larijani B, Nasiri S, Tavangar SM. Liquid biopsy as a minimally invasive source of thyroid cancer genetic and epigenetic alterations. Int J Mol Cell Med. 2019;8(Suppl 1):19-29.
(74) Rappa G, Puglisi C, Santos MF, Forte S, Memeo L, Lorico A. Extracellular vesicles from thyroid carcinoma: the new frontier of liquid biopsy. Int J Mol Sci. 2019;20(5).
(75) Liang M, Yu S, Tang S, Bai L, Cheng J, Gu Y, et al. A panel of plasma exosomal miRNAs as potential biomarkers for differential diagnosis of thyroid nodules. Front Genet. 2020;11:449.
(76) Mao Y, Zhang F, He L, Luo F, Li L, Huo Y, et al. Added value of circulating miRNA expression profiling to sonographic TI-RADS classification in the diagnosis of thyroid nodules. Exp Ther Med. 2020;20(2):1589-95.
(77) Galimzianova A, Siebert SM, Kamaya A, Rubin DL, Desser TS. Quantitative framework for risk stratification of thyroid nodules with ultrasound: a step toward automated triage of thyroid cancer. AJR Am J Roentgenol. 2020;214(4):885-92.
(78) DeHoog RJ, Zhang J, Alore E, Lin JQ, Yu W, Woody S, et al. Preoperative metabolic classification of thyroid nodules using mass spectrometry imaging of fine-needle aspiration biopsies. Proc Natl Acad Sci U S A. 2019;116(43):21401-8.
(79) Colakoglu B, Alis D, Yergin M. Diagnostic value of machine learning-based quantitative texture analysis in differentiating benign and malignant thyroid nodules. J Oncol. 2019;2019:6328329.
(80) Ravella L, Lopez J, Descotes F, Giai J, Lapras V, Denier ML, et al. Cytological features and nuclear scores: diagnostic tools in preoperative fine needle aspiration of indeterminate thyroid nodules with RAS or BRAF K601E mutations? Cytopathology. 2021;32(1):37-44.
(81) Cooper DS, Doherty G, Haugen B, Kloos RT. Management guidelines for patients with thyroid nodules and differentiated thyroid cancer. Thyroid. 2009;19(11):1167-214.
(82) National Institute for Health and Care Excellence. Process and methods guides. Appendix I: Quality appraisal checklist–economic evaluations [Internet]. London: The Institute; 2012 [cited 2016 Jan]. Available from: https://www.nice.org.uk/process/pmg4/chapter/appendix-i-quality-appraisal-checklist-economic-evaluations
(83) Balentine CJ, Vanness DJ, Schneider DF. Cost-effectiveness of lobectomy versus genetic testing (Afirma) for indeterminate thyroid nodules: considering the costs of surveillance. Surgery. 2018;163(1):88-96.

(84) Labourier E. Utility and cost-effectiveness of molecular testing in thyroid nodules with indeterminate cytology. Clin Endocrinol (Oxf). 2016;85(4):624-31.
(85) Vriens D, Adang EM, Netea-Maier RT, Smit JW, de Wilt JH, Oyen WJ, et al. Cost-effectiveness of FDG-PET/CT for cytologically indeterminate thyroid nodules: a decision analytic approach. J Clin Endocrinol Metab. 2014;99(9):3263-74.
(86) Yip L, Farris C, Kabaker AS, Hodak SP, Nikiforova MN, McCoy KL, et al. Cost impact of molecular testing for indeterminate thyroid nodule fine-needle aspiration biopsies. J Clin Endocrinol Metab. 2012;97(6):1905-12.
(87) Zanocco KA, Wang MM, Yeh MW, Livhits MJ. Selective use of molecular testing based on sonographic features of cytologically indeterminate thyroid nodules: a decision analysis. World J Surg. 2020;44(2):393-401.
(88) Fazeli SR, Zehr B, Amraei R, Toraldo G, Guan H, Kindelberger D, et al. ThyroSeq v2 testing: impact on cytologic diagnosis, management, and cost of care in patients with thyroid nodule. Thyroid. 2020;19:19.
(89) Rivas AM, Nassar A, Zhang J, Casler JD, Chindris AM, Smallridge R, et al. ThyroSeq v2.0 molecular testing: a cost-effective approach for the evaluation of indeterminate thyroid nodules. Endocrine Practice. 2018;24(9):780-8.
(90) Abeykoon JP, Mueller L, Dong F, Chintakuntlawar AV, Paludo J, Mortada R. The effect of implementing gene expression classifier on outcomes of thyroid nodules with indeterminate cytology. Horm Cancer. 2016;7(4):272-8.
(91) Shapiro S, Pharaon M, Kellermeyer B. Cost-effectiveness of gene expression classifier testing of indeterminate thyroid nodules utilizing a real cohort comparator. Otolaryngol Head Neck Surg. 2017;157(4):596-601.
(92) Wu JX, Lam R, Levin M, Rao J, Sullivan PS, Yeh MW. Effect of malignancy rates on cost-effectiveness of routine gene expression classifier testing for indeterminate thyroid nodules. Surgery. 2016;159(1):118-26.
(93) Lee L, How J, Tabah RJ, Mitmaker EJ. Cost-effectiveness of molecular testing for thyroid nodules with atypia of undetermined significance cytology. J Clin Endocrinol Metab. 2014;99(8):2674-82.
(94) Li H, Robinson KA, Anton B, Saldanha IJ, Ladenson PW. Cost-effectiveness of a novel molecular test for cytologically indeterminate thyroid nodules. J Clin Endocrinol Metab. 2011;96(11):E1719-26.
(95) Nicholson KJ, Roberts MS, McCoy KL, Carty SE, Yip L. Molecular testing versus diagnostic lobectomy in Bethesda III/IV thyroid nodules: a cost-effectiveness analysis. Thyroid. 2019;29(9):1237-43.
(96) Ministry of Health, Ministry of Long-Term Care. Health data branch web portal: Ontario case costing [Internet]. Toronto (ON): Queen's Printer for Ontario. 2021 [cited 2021 Oct 18]. Available from: https://hsim.health.gov.on.ca/HDBPortal
(97) Husereau D, Drummond M, Petrou S, Carswell C, Moher D, Greenberg D, et al. Consolidated Health Economic Evaluation Reporting Standards (CHEERS)--explanation and elaboration: a report of the ISPOR Health Economic Evaluation Publication Guidelines Good Reporting Practices Task Force. Value Health. 2013;16(2):231-50.
(98) Canadian Agency for Drugs and Technologies in Health. Guidelines for the economic evaluation of health technologies: Canada. 4th ed. Ottawa (ON): The Agency; 2017. p. 76.
(99) Cancer Care Ontario. Differentiated thyroid cancer diagnosis pathway map. Version 2019.09 [Internet]. Toronto (ON): Cancer Care Ontario; 2019 [cited 2021 June 10]. Available from: https://www.cancercareontario.ca/sites/ccocancercare/files/assets/CCOThyroidDiagnosisPathway.pdf

(100) Statistics Canada. Life tables, Canada, provinces and territories [Internet]. 2021 [cited 2021 Oct 18]. Available from: https://www150.statcan.gc.ca/n1/en/catalogue/84-537-X
(101) Al-Qurayshi Z, Farag M, Shama MA, Ibraheem K, Randolph GW, Kandil E. Total thyroidectomy versus lobectomy in small nodules suspicious for papillary thyroid cancer: cost-effectiveness analysis. Laryngoscope. 2020;130(12):2922-6.
(102) Noel CW, Griffiths R, Siu J, Forner D, Urbach D, Freeman J, et al. A population-based analysis of outpatient thyroidectomy: safe and under-utilized. Laryngoscope. 2021;131:2625-33.
(103) Gómez-Ramírez J, Sitges-Serra A, Moreno-Llorente P, Zambudio AR, Ortega-Serrano J, Rodríguez MT, et al. Mortality after thyroid surgery, insignificant or still an issue? Langenbecks Arch Surg. 2015;400(4):517-22.
(104) Guertin JR, Feeny D, Tarride JE. Age- and sex-specific Canadian utility norms, based on the 2013-2014 Canadian Community Health Survey. CMAJ. 2018;190(6):E155-e61.
(105) Ministry of Health. Schedule of benefits: physician services under the Health Insurance Act. Toronto (ON): Queen's Printer for Ontario; 2021.
(106) Ministry of Health. Schedule of benefits for laboratory services. Toronto (ON): Queen's Printer for Ontario; 2021.
(107) Statistics Canada. Consumer Price Index: table 326-0020 [Internet]. Ottawa (ON): Statistics Canada; 2018 [cited 2021 Oct 18]. Available from: http://www5.statcan.gc.ca/cansim/a26?lang=eng&id=3260020
(108) Canadian Institute for Health Information. Cost of a standard hospital stay [Internet]. Ottawa (ON): The Institute; 2020 [cited 2021 Apr 20]. Available from: https://yourhealthsystem.cihi.ca/hsp/inbrief?lang=en#!/indicators/015/cost-of-a-standard-hospital-stay-cshs/;mapC1;mapLevel2;provinceC9001;/
(109) Canadian Institute for Health Information. Hospital spending: focus on the emergency department [Internet]. Ottawa (ON): The Institute; 2020 [cited 2021 Apr 20]. Available from: https://www.cihi.ca/sites/default/files/document/hospital-spending-highlights-2020-en.pdf
(110) Tam S, Sun H, Sarma S, Siu J, Fung K, Sowerby L. Medialization thyroplasty versus injection laryngoplasty: a cost minimization analysis. J Otolaryngol Head Neck Surg. 2017;46(1):14.
(111) James BC, Aschebrook-Kilfoy B, White MG, Applewhite MK, Kaplan SP, Angelos P, et al. Quality of life in thyroid cancer-assessment of physician perceptions. J Surg Res. 2018;226:94-9.
(112) Ferris RL, Baloch Z, Bernet V, Chen A, Fahey TJ, 3rd, Ganly I, et al. American Thyroid Association statement on surgical application of molecular profiling for thyroid nodules: current impact on perioperative decision making. Thyroid. 2015;25(7):760-8.
(113) Sacks WL, Bose S, Zumsteg ZS, Wong R, Shiao SL, Braunstein GD, et al. Impact of Afirma gene expression classifier on cytopathology diagnosis and rate of thyroidectomy. Cancer Cytopathol. 2016;124(10):722-8.
(114) Barham L. Public and patient involvement at the UK National Institute for Health and Clinical Excellence. Patient. 2011;4(1):1-10.
(115) Messina J, Grainger DL. A pilot study to identify areas for further improvements in patient and public involvement in health technology assessments for medicines. Patient. 2012;5(3):199-211.
(116) Ontario Health Technology Advisory Committee Public Engagement Subcommittee. Public engagement for health technology assessment at Health Quality Ontario—final report from the Ontario Health Technology Advisory Committee Public Engagement Subcommittee [Internet]. Toronto (ON): Queen's Printer for Ontario; 2015 Apr [cited 2018 Apr 30]. Available from: http://www.hqontario.ca/Portals/0/documents/evidence/special-reports/report-subcommittee-20150407-en.pdf

(117) Lee DJ, Xu JJ, Brown DH, Gilbert RW, Gullane PJ, Irish JC, et al. Determining patient preferences for indeterminate thyroid nodules: observation, surgery or molecular tests. World J Surg. 2017;41(6):1513-20.
(118) Wong CW, Schumm MA, Zhu CY, Tseng CH, Arasu A, Han J, et al. Quality of life following molecular marker testing for indeterminate thyroid nodules. Endocrine Practice. 2020;14:14.
(119) Kvale S. Interviews: an introduction to qualitative research interviewing. Thousand Oaks (CA): Sage; 1996.
(120) Kuzel AJ. Sampling in qualitative inquiry. In: Miller WL, Crabtree BF, editors. Doing qualitative research. Thousand Oaks (CA): Sage; 1999. p. 33–45.
(121) Morse J. Emerging from the data: cognitive processes of analysis in qualitative research. In: Morse J, editor. Critical issues in qualitative research methods. Thousand Oaks (CA): Sage; 1994. p. 23-41.
(122) Patton MQ. Qualitative research and evaluation methods. 3rd ed. Thousand Oaks (CA): Sage; 2002.
(123) Strauss AL, Corbin JM. Basics of qualitative research: techniques and procedures of developing a grounded theory. 2nd ed. Thousand Oaks (CA): Sage; 1998.
(124) Health Technology Assessment International. Introduction to health technology assessment [Internet]. Edmonton (AB): Health Technology Assessment International; 2015 [cited 2018 Apr 30]. Available from: http://www.htai.org/fileadmin/HTAi_Files/ISG/PatientInvolvement/v2_files/Resource/PCISG-Resource-Intro_to_HTA__KFacey__Jun13.pdf
(125) Strauss AL, Corbin JM. Grounded theory research: procedures, canons, and evaluative criteria. Qual Sociol. 1990;13(1):3-21.
(126) Strauss AL, Corbin JM. Grounded theory methodology: an overview. In: Denzin NK, Lincoln YS, editors. Handbook of qualitative research. Thousand Oaks (CA): Sage; 1994. p. 273-85.
(127) NVivo qualitative data analysis software. QSR International. Doncaster, Victoria (Australia). Available at: https://www.qsrinternational.com/nvivo/home.
(128) Borowczyk M, Szczepanek-Parulska E, Olejarz M, Wieckowska B, Verburg FA, Debicki S, et al. Evaluation of 167 gene expression classifier (GEC) and ThyroSeq v2 diagnostic accuracy in the preoperative assessment of indeterminate thyroid nodules: bivariate/HROC meta-analysis. Endocr Pathol. 2019;30(1):8-15.
(129) Borowczyk M, Szczepanek-Parulska E, Olejarz M, Wieckowska B, Verburg FA, Debicki S, et al. Correction to: Evaluation of 167 gene expression classifier (GEC) and ThyroSeq v2 diagnostic accuracy in the preoperative assessment of indeterminate thyroid nodules: bivariate/HROC meta-analysis. Endocr Pathol. 2019;30(1):16.
(130) Liu Y, Pan B, Xu L, Fang D, Ma X, Lu H. The diagnostic performance of Afirma gene expression classifier for the indeterminate thyroid nodules: a meta-analysis. BioMed Res Int. 2019;2019:7150527.
(131) Muzza M, Colombo C, Pogliaghi G, Karapanou O, Fugazzola L. Molecular markers for the classification of cytologically indeterminate thyroid nodules. J Endocrinol Invest. 2020;43(6):703-16.
(132) Ngo HTT, Nguyen TPX, Vu TH, Jung CK, Hassell L, Kakudo K, et al. Impact of molecular testing on the management of indeterminate thyroid nodules among Western and Asian countries: a systematic review and meta-analysis. Endocr Pathol. 2020;7:7.
(133) Valderrabano P, Hallanger-Johnson JE, Thapa R, Wang X, McIver B. Comparison of postmarketing findings vs the initial clinical validation findings of a thyroid nodule gene expression classifier: a systematic review and meta-analysis. JAMA Otolaryngol Head Neck Surg. 2019;18:18.

(134) Vuong HG, Nguyen TPX, Hassell LA, Jung CK. Diagnostic performances of the Afirma gene sequencing classifier in comparison with the gene expression classifier: a meta-analysis. Cancer Cytopathol. 2020;29:29.
(135) GRADE handbook [Internet]. 2013 [updated 2013; cited 2021 Mar 22]. Available from: https://gdt.gradepro.org/app/handbook/handbook.html#h.xivvyiu1pr3v
(136) Singer J, Hanna JW, Visaria J, Gu T, McCoy M, Kloos RT. Impact of a gene expression classifier on the long-term management of patients with cytologically indeterminate thyroid nodules. Curr Med Res Opin. 2016;32(7):1225-32.
(137) Najafzadeh M, Marra CA, Lynd LD, Wiseman SM. Cost-effectiveness of using a molecular diagnostic test to improve preoperative diagnosis of thyroid cancer. Value Health. 2012;15(8):1005-13.
(138) Chudova D, Wilde JI, Wang ET, Wang H, al e. Molecular classification of thyroid nodules using high-dimensionality genomic data. J Clin Endocrinol Metab. 2010;95(12):5296-304.

About Us

Ontario Health is an agency of the Government of Ontario. Our mandate is to connect and coordinate our province's health care system in ways that have not been done before to help ensure that Ontarians receive the best possible care. We work to support better health outcomes, patient experiences, provider experiences and value for money spent.

For more information about Ontario Health, visit ontariohealth.ca.

About the Ontario Health Technology Advisory Committee

How to Obtain Reports From the Ontario Health Technology Assessment Series

Disclaimer

Ontario Health
130 Bloor Street West, 10th Floor
Toronto, Ontario
M5S 1N5
Toll Free: 1-877-280-8538
TTY: 1-800-855-0511
Email: oh-hqo_hta@ontariohealth.ca
www.hqontario.ca

ISSN 1915-7398 (online)
ISBN 978-1-4868-5888-0 (PDF)

Ontario Health