

# Molecular Testing Versus Diagnostic Lobectomy in Bethesda III/IV Thyroid Nodules: A Cost-Effectiveness Analysis

Kristina J. Nicholson,<sup>1</sup> Mark S. Roberts,<sup>2,3</sup> Kelly L. McCoy,<sup>1</sup> Sally E. Carty,<sup>1</sup> and Linwah Yip<sup>1</sup>

**Background:** Molecular tests (MT) using gene expression and/or mutational analysis have been developed to reduce the need for diagnostic surgery for indeterminate (Bethesda III/IV) thyroid nodules. Prior cost-effectiveness studies have shown mixed results but none has included the recent and more comprehensive versions of the two commonly utilized MT. The aim of this study is to compare the cost-effectiveness of diagnostic lobectomy (DL), the Afirma Gene Sequencing Classifier (GSC), and ThyroSeq version 3 (TSv3).

**Methods:** A decision tree from the payer perspective was created using a base case of a 40-year-old euthyroid woman with a solitary 2 cm Bethesda III or IV thyroid nodule. In this model, all patients in the DL arm had lobectomy, which was also performed for patients with positive MT, while those with negative MT underwent 20 years of surveillance. The outcome was a correct diagnosis, defined as malignant histology after DL or 20 years of nodule stability after negative MT. Costs were obtained from the Centers for Medicare & Medicaid Services (CMS) data and existing literature, and probabilities were obtained from the literature. Sensitivity analysis was performed for costs, pretest probability of malignancy, and performance parameters.

**Results:** The cost per correct diagnosis was \$14,277 for TSv3, \$17,873 for GSC, and \$38,408 for DL. TSv3 was preferred over both GSC and DL. One-way sensitivity analysis between TSv3 and GSC demonstrated that the results were robust to variations in cost, cancer prevalence, and length of surveillance. In the two-way sensitivity analysis, TSv3 was preferred over GSC at all considered test costs, and in probabilistic sensitivity analysis, TSv3 was the preferred management strategy in 68.5% of cases.

**Conclusions:** In hypothetical modeling to determine whether surgery versus MT is optimal for indeterminate (Bethesda III/IV) nodules, either of the major MT was considerably more cost-effective than DL, although TSv3 was more likely to be cost-effective than GSC. Use of either MT adjunct should be strongly considered in the absence of other indications for thyroidectomy.

**Keywords:** thyroid nodule, molecular testing, cost-effectiveness, Bethesda III/IV nodule, diagnostic lobectomy

## Introduction

**B**EFORE THE DEVELOPMENT of molecular tests (MT), thyroid nodules with fine needle aspiration biopsy (FNAB) results classified in categories III or IV of the Bethesda System for Reporting Thyroid Cytopathology (1) typically required additional evaluation, including diagnostic surgery for definitive diagnosis (2,3), contributing significantly to the increasing number of thyroidectomies that are performed every year in the United States (4).

As knowledge has grown about the genetic mutations and molecular alterations of thyroid cancer (5), MT have been

developed that reliably identify lesions at risk of malignancy based on gene sequencing and mRNA expression analysis, and are now used to obviate the need for diagnostic surgery. However, in addition to concerns about MT effectiveness and utility outside of the initial validation setting, the cost of such testing is a relevant consideration, especially with the shifting reimbursement structure in the United States (6).

Prior cost analyses of thyroid MT have yielded a variety of results, with some tests demonstrating cost savings and improved quality of life (7–11), while others have shown the opposite (12–14). Factors that may contribute to poor MT cost-effectiveness include test cost and test accuracy (12,13).

<sup>1</sup>Division of Endocrine Surgery, Department of Surgery, University of Pittsburgh, Pittsburgh, Pennsylvania.  
Division of General Internal Medicine, Departments of <sup>2</sup>Medicine and <sup>3</sup>Health Policy and Management, University of Pittsburgh, Pittsburgh, Pennsylvania.

Presented as an oral abstract at the 88th Meeting of the American Thyroid Association, Washington, District of Columbia, October 6, 2018.

The sensitivity and specificity of the two most commonly used MT for thyroid nodules have recently improved. The Afirma (Veracyte, Inc., South San Francisco, CA) Gene Expression Classifier (GEC) has been expanded to the new Gene Sequencing Classifier (GSC) and in a clinical validation study, the sensitivity was 91% and specificity improved to 68% (15). ThyroSeq version 3 (TSv3; CBLPath, Rye Brook, NY) now includes comprehensive gene expression and copy number variation data for 112 thyroid cancer-related genes (16), and in a recent multi-institutional clinical validation study demonstrated a sensitivity and a specificity of 94% and 82%, respectively (17).

For cytologically indeterminate thyroid lesions, earlier MT versions were reportedly cost effective in clinical management, using extant versions of the American Thyroid Association (ATA) guidelines (7,8). However, no study has compared the cost efficacy of diagnostic lobectomy (DL) with the newest MT versions, included recommendations for extent of thyroidectomy under the 2015 ATA guidelines, or compared the two MT with each other and with DL. Our study aim was to directly assess the cost-effectiveness of these three management strategies (DL, GSC, TSv3) for Bethesda III and IV thyroid nodules using the current clinical management algorithms of the 2015 ATA guidelines (18) and optimally matched management and diagnosis as the outcome measure.

**Materials and Methods**

*Decision model*

A decision tree representing common management strategies for a base case of a 40-year-old euthyroid woman with a solitary 2 cm Bethesda III or IV (BIII/IV, i.e., cytologically indeterminate) thyroid nodule was constructed in TreeAge Pro 2018 (TreeAge Software, Inc., Williamstown, MA). The hypothetical model compared DL, TSv3, and GSC (Fig. 1) (15–19). Since no human research was being carried out, a University of Pittsburgh IRB waiver was obtained.

Briefly, all patients in the DL arm had thyroid lobectomy. In addition, thyroid lobectomy was performed for patients with positive TSv3 or suspicious GSC results, or when MT failed to produce an interpretable result. Negative GSC or

TSv3 results led to a cohort Markov model for surveillance with a one-year cycle length for a base case time horizon of 20 years. During each annual surveillance cycle, patients had a follow-up clinic visit and neck/thyroid ultrasound, at which time they could remain well with a stable nodule, have thyroid cancer detected and undergo lobectomy, or die from nonthyroid causes due to background mortality (Fig. 1 and Table 1) (20).

The base case cancer prevalence for a BIII/IV nodule was 25% (5–50%) (18). The probability of thyroid cancer detection during surveillance was calculated based upon the probability of false negative results for both tests in the literature, and divided over years of surveillance (15,17). As there is currently no literature defining the adequate length of follow-up after negative MT, based on expert opinion a surveillance length of 20 years was used in the base case analysis.

*Costs*

Costs were considered from a U.S. payer perspective, which includes costs incurred by third-party payers (21), for a cost year of 2018. The initial costs incurred during nodule evaluation (clinic visit, laboratory evaluation, ultrasound, and biopsy, which resulted in a BIII/IV diagnosis) were not included as all three strategies would have equivalent costs.

Each arm included specific costs related to the management strategy and related surveillance. Costs in the DL arm included the cost of lobectomy and the lifetime cost of lobectomy-associated complications, incorporated as a distributed cost by multiplying the rates and cost of each major complication type (hematoma, hypothyroidism, and vocal cord dysfunction) and applying the sum of the products to each instance of lobectomy (Table 1) (5,11). Initial costs in the MT arms included cost of a second biopsy, to account for the possibility of repeat FNA for a BIII nodule and/or need for repeat biopsy for material for MT, in addition to the cost of either TSv3 or GSC.

In the event of positive MT or when MT failed to produce an interpretable result, costs of lobectomy and complications were included. In the event of negative MT, costs for each yearly Markov cycle included the cost of a clinic visit and ultrasound. When cancer was detected during surveillance after negative MT, additional costs incurred during

**FIG. 1.** Decision tree used to compare rates of correct diagnosis for DL, TSv3, and GSC (16). 1: effectiveness of optimal management, 0: effectiveness of nonoptimal management. In the Markov model, remaining in the “Well” category at the conclusion of the model results in an effectiveness value of 1, while transitioning to the “Detect cancer” or “Dead” states at any time results in an effectiveness of 0. DL, diagnostic lobectomy; GSC, Gene Sequencing Classifier; TSv3, ThyroSeq version 3.

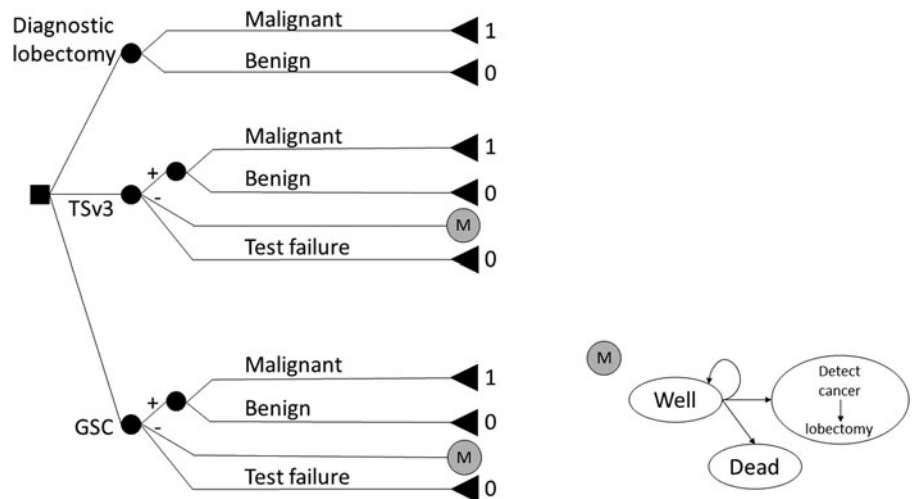


TABLE 1. MODEL INPUTS FOR BASE CASE ANALYSIS

Variable	Value (range)	Source
Probabilities and test parameters		
Cancer prevalence	0.25 (0.05–0.5)	(15)
MT failure	0.10 (0.05–0.15)	(12,14)
TSv3 sensitivity	0.94 (0.86–0.98)	(14)
TSv3 specificity	0.82 (0.75–0.87)	(14)
GSC sensitivity	0.91 (0.79–0.98)	(12)
GSC specificity	0.68 (0.60–0.76)	(12)
Postoperative hematoma	0.008 (0.008–0.017)	(16)
Postoperative hypothyroidism	0.25 (0.15–0.35)	(15)
Postoperative transient RLN palsy	0.036 (0.01–0.07)	(15)
Postoperative permanent RLN palsy	0.021 (0.01–0.035)	(15)
Surveillance of MT-negative nodule (years)	20 (5–20)	<sup>a</sup>
Costs (\$)		
Biopsy and cytology interpretation	142 (114–170)	(18)
Follow-up clinic visit	106 (85–127)	(16)
Thyroid ultrasound	120 (96–144)	(18)
GSC testing	3600 (2560–7680)	(7,18,19)
TSv3 testing	3600 (2880–4867)	(18,19)
DL	9520 (6235–12,471)	(7,16)
Complications of lobectomy		
Hematoma	6336 (5069–7603)	(16)
Hypothyroidism	1852 (1482–2222)	(7)
Transient RLN palsy	2365 (1892–2838)	(7)
Permanent RLN palsy	7044 (5635–8453)	(7)

Ranges for sensitivity analysis provided in parentheses.

<sup>a</sup>Expert opinion.

DL, diagnostic lobectomy; GSC, Gene Sequencing Classifier; MT-negative, molecular test negative; TSv3, ThyroSeq version 3.

that cycle included the cost of biopsy and cytology, and the cost of lobectomy and associated complications. Costs were obtained from the literature as cited, and the Centers for Medicare & Medicaid Services (CMS) 2018 reimbursement schedule (Table 1) (22).

In the base case, \$3600 was used for the cost of both types of MT, per 2018 Medicare reimbursement (22). For sensitivity analysis, previously published costs for the tests were included and varied by  $\pm 20\%$  to account for a wide range of possible future test costs (11,22,23). Costs that were obtained from literature published >1 year ago were adjusted for inflation to June 2018 costs using the Consumer Price Index inflation calculator (3) and were rounded to the nearest dollar. Future costs were discounted at a standard 3% rate (21).

### Outcomes

The outcome of interest was optimal nodule management defined as surgery for nodules with malignancy (including noninvasive follicular thyroid neoplasm with papillary-like nuclear features [NIFTP]) on final histology and surveillance for those that are truly benign. In the context of MT, this can be summarized as accurately predicting the correct diagnosis,

and so it will be called herein. In terms of effectiveness, this was considered a binary outcome, such that a correct diagnosis was assigned an effectiveness of 1, while an incorrect diagnosis was assigned an effectiveness of 0. Scenarios in which surgery was performed for a histologically benign nodule were considered incorrect diagnoses, as were scenarios in which a missed malignancy was diagnosed on lobectomy during surveillance after negative MT.

As histologic evaluation is required for NIFTP diagnosis, surgery is required and when performed was considered the correct management strategy. Technical failure of MT was assigned an effectiveness of 0 regardless of histologic outcome. Patients who were lost to follow-up during surveillance due to background mortality were also assigned an effectiveness score of 0. Patients who die of unrelated causes during surveillance are likely to be true negatives, but because they did not complete the assigned rigorous follow-up, we did not want to overestimate correct nodule diagnosis. The overall effectiveness for each strategy was represented as the probability of correct diagnosis for each management strategy. Cost and effectiveness of the base case were summarized as the cost per correct diagnosis.

### Sensitivity analysis

One-way sensitivity analysis was performed for costs of DL, TSv3, and GSC, cancer prevalence in nodules with BIII/BIV FNAB results, performance parameters for TSv3 and GSC, MT failure rate, and length of MT-negative nodule surveillance (Table 1). Ranges were obtained using values in the literature or by varying the base case by 20% when literature was not available. Two-way sensitivity analysis was performed to determine the MT cost at which the preferred treatment strategy would change. Probabilistic sensitivity analysis was performed using beta distributions for probabilities and test parameters, gamma distributions for costs, and a Poisson distribution for nodule surveillance time in years.

### Results

DL had the lowest average expected cost per base case, at \$9602, compared with \$10,451 and \$11,385 for TSv3 and GSC, respectively (Table 2). However, TSv3 had the highest probability of a correct diagnosis at 0.732 compared with 0.637 for GSC and 0.250 for DL. In other words, when TSv3 was used, the likelihood of obtaining a correct diagnosis was 73.2%. Therefore, the cost per correct diagnosis was lowest for TSv3 at \$14,277 compared with \$17,873 for GSC, and \$38,408 for DL (Table 2).

One-way sensitivity analysis was performed to test model assumptions (Fig. 2). Within the tested ranges in Table 1, TSv3 had a robustly lower cost per correct diagnosis than GSC. When tested ranges were broadened beyond the predicted variation in Table 1, extreme thresholds outside likely real-world values were identified. In this latter scenario, GSC became less costly per diagnosis when TSv3 sensitivity decreased to 0.76 (base case 0.94) or specificity decreased to 0.66 (base case 0.82), or when GSC specificity increased to 0.86 (base case 0.68).

When the cost of GSC was held constant at the current CMS reimbursement rate of \$3600 (22), TSv3 became more costly if the reimbursement was >\$6218. Similarly, when the cost of TSv3 was held constant at \$3600, GSC became less

TABLE 2. RESULTS: COST AND EFFECTIVENESS OF DIAGNOSTIC LOBECTOMY, THYROSEQ VERSION 3, AND GENE SEQUENCING CLASSIFIER

Model strategy	Cost/case (\$)	Probability of correct diagnosis	Cost/correct diagnosis (\$)
DL	9602	0.250	38,408
TSv3	10,451	0.732	14,277
GSC	11,385	0.637	17,873

costly if cost was decreased to <\$1318. Altogether, the model was robust to broad variations in the cost of TSv3, GSC, and DL, sensitivity and specificity of both TSv3 and GSC, length of surveillance after negative MT, and cancer prevalence in BIII/IV nodules.

Two-way sensitivity analysis considering simultaneous variations in cost of TSv3 and GSC demonstrated that TSv3 remained the preferred strategy over GSC across all tested ranges in cost (Table 1). When the model utilized the most recently published costs of GSC and TSv3 (\$6400 and \$4056, respectively) (23), the cost per correct diagnosis was \$14,900 for TSv3 and \$22,268 for GSC. In no range of tested cost variations did DL become the preferred strategy over MT.

On probabilistic sensitivity analysis, the model was stable at a value of 10,000 iterations. TSv3 was the preferred strategy in 68.5% of cases, while GSC and DL were preferred 25.0% and 6.5% of the time, respectively (Fig. 3).

## Discussion

In this theoretical model, we (i) used current management algorithms for Bethesda III/IV nodules (18) and (ii) utilized the most recent costs and probabilities from the newest validation data for both GSC and TSv3 (15,17). In addition, we directly compared DL with both GSC and TSv3. Despite the choice of a long MT surveillance interval (20 years) that would potentially skew model results to favor DL, the first major finding of the study was that although DL had the lowest

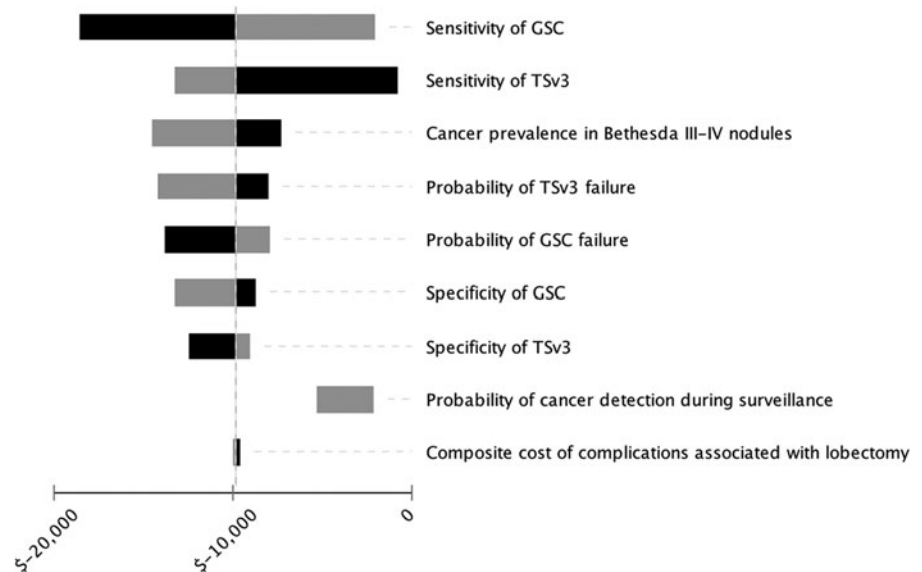
cost per case, use of either MT was found to have significantly lower cost per correct diagnosis than DL. The cost savings observed for both MT strategies are the direct result of avoiding unnecessary diagnostic surgery for what proves later to be a histologically benign thyroid nodule. These results have significant cost and safety implications for patients, as well as a meaningful potential economic impact in the United States.

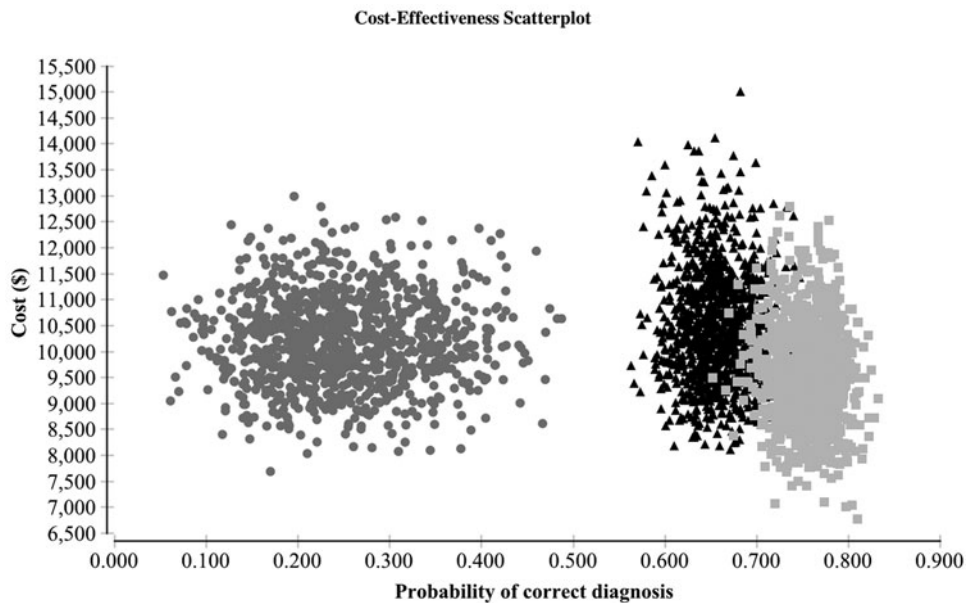
MT cost efficacy compared with DL has been previously studied in some detail, but not using the most recently available tests. In the first cost utility study of MT, evaluating GEC compared with DL, Li *et al.* demonstrated both cost savings and a modest gain in quality-adjusted life years (QALYs) in the MT arm of a hypothetical model (7). The cost savings persisted up to an MT cost of \$4600, and MT was cost saving in 92.5% of tested cases on probabilistic sensitivity analysis. Similarly, Labourier demonstrated cost savings likely resulting from a 32% reduction in unnecessary surgery in a 2016 theoretical analysis of a different MT called ThyGenX/ThyraMIR (11).

In a hypothetical cost study evaluating the seven-gene panel (on which TSv3 was originally based) and using clinical management as per the existing 2009 ATA guideline recommendation for total thyroidectomy for cancer >1 cm (3), cost savings were observed due to reductions in two-stage thyroidectomy (8). A recent study of TSv2 using patient-level cost data demonstrated a decrease in overall costs by avoiding unnecessary surgery, however, the design did not account for any follow-up of MT-negative nodules (10) or consider current management recommendations that include lobectomy as adequate treatment for low-grade thyroid malignancies (18).

Interestingly, not all studies of MT compared with DL have demonstrated cost savings. For example, in 2017, Shapiro *et al.* integrated real patient cost data from a small cohort of 96 patients with indeterminate nodules and found only a 13.1% reduction in necessary surgeries with GEC, and increased cost in the GEC cohort (13). In this study, the indications and criteria for GEC testing were not clear and management efficacy was assessed retrospectively. Similarly, a recent cost utility analysis by Balentine *et al.* showed

FIG. 2. Tornado diagram: one-way sensitivity analysis of incremental cost-effectiveness of TSv3 over GSC. Black bars represent higher parameter values and gray bars represent lower parameter values. Values on x-axis represent cost savings per correct diagnosis of TSv3 over GSC.





**FIG. 3.** Probabilistic sensitivity analysis: cost-effectiveness scatter plot demonstrating cost and probability of correct diagnosis over 10,000 iterations with simultaneous variation of input values. Circle: diagnostic lobectomy; triangle: GSC; square: TSv3.

that compared with DL, GEC was associated with both a modest decrease in QALYs and an increase in cost, which was attributed to surveillance for GEC-negative nodules (14).

In addition, and different than the present model, many patients with positive MT went on to have total thyroidectomy rather than lobectomy, incurring additional costs and risks. Here, because the current ATA guidelines recommend DL or total thyroidectomy as oncologically equivalent definitive surgical options for 1–4 cm confined, differentiated thyroid cancers, which are typically the cancers diagnosed after resection of BIII/BIV nodules, we utilized DL as the primary surgical strategy. Completion thyroidectomy may be needed for a proportion of histologic cancers; however, this proportion would likely be equivalent in all three management strategies.

Several prior cost-effectiveness studies of MT have focused on cost per QALY gained and have shown minimal differences in QALYs, regardless of the management strategy (7,9,12,14). This is likely because thyroid surgery is relatively safe and differentiated thyroid cancers are generally indolent with missed cancers rarely causing significant morbidity or mortality. Thus, in the present study design, we focused on an objective and time-independent outcome, which is an optimally matched therapy and diagnosis, since the purpose of obtaining additional testing is to arrive at an accurate diagnosis and direct appropriate management. We also believe that this effectiveness measure represents a more practical approach that provides a clearer picture of how adjunct tests such as MT can affect health care expenditures.

Another finding was that when GSC and TSv3 were compared, TSv3 was significantly preferred, with both lower cost and a higher likelihood of directing the correct clinical management according to histologic outcome. Over a broad range of variation in values, TSv3 was consistently the least costly strategy. Validation studies of GEC and earlier versions of ThyroSeq have demonstrated mixed results in diagnostic accuracy (24–34); the current model preference for TSv3 over GSC was robust with different results obtained only when TSv3 and GSC specificities were well outside the

previously reported 95% confidence interval thresholds (of 0.66 and 0.86, respectively) (15,17). On expanded sensitivity analysis, test preference did change at a low cost of GSC or high cost of TSv3, as expected.

Cancer prevalence has been widely highlighted as a potential source for heterogeneity in MT accuracy. In the current model, TSv3 and GSC were preferred to DL over a wide (5–50%) cancer prevalence range considered for indeterminate nodules, suggesting that either type of MT is more cost effective than routine DL even when cancer prevalence in BIII/BIV nodules is as high as 50%. As the national reimbursement landscape shifts increasingly toward bundled payments and diagnosis-related groups, careful consideration of economical and effective management strategies is becoming increasingly important, and reconsideration of costly reflexive DL for BIII/IV nodules may be worthwhile (6).

This study has several limitations. First, the diagnostic accuracy of MT has the potential to be decreased by the introduction of new terminology to describe NIFTP (35). Because NIFTP is currently considered a lesion requiring surgical resection for diagnosis and management, considering it to be malignant for the purposes of this analysis was the most accurate way to analyze the cost efficacy of correct management. If in the future, NIFTP lesions should become identifiable preoperatively and not require resection, the diagnostic accuracy of MT and the results of our model would be altered. Regardless, the cost per correct diagnosis for either TSv3 or GSC would need to increase by several fold before DL would become the preferred strategy.

Furthermore, the study is theoretical, that is, the costs and efficacies utilized were evidence-based estimated values that may not represent the costs and probabilities of any one patient, clinical scenario, or geographical/institutional setting. The model does not account for variations in practice such as cytology rereview or repeat FNAB of nodules before proceeding to MT or DL; to minimize the impact of this variability and remain generalizable, it included costs incurred after these measures, when surgery becomes the next consideration in the diagnostic algorithm. Next, the initial

performance parameters of GSC and TSv3 were reported in single multicenter studies and both require further independent validation (15–17); in future study, it is possible that test performance parameters may lie outside those used here on sensitivity analysis.

In addition, this study considered cost only up until the time of definitive diagnosis. Further costs, which are likely equivalent between the three management strategies, may be incurred with postlobectomy treatment of cancer and were not accounted for, including completion thyroidectomy and possibly radioactive iodine administration. Given these considerations and the binary “correct diagnosis” outcome measure, which has no known threshold for cost-effectiveness (i.e., \$100,000/QALY), these results should be interpreted primarily as a relative cost for each strategy, rather than as an absolute cost per diagnosis.

Finally, we used a base case of a 40-year-old woman with a 2 cm Bethesda III/IV thyroid nodule with no other concerning symptomatic or sonographic features. This situation is not necessarily representative of all cases, and certainly, individual factors must be considered before even ordering MT; that is, patients may present with compressive symptoms warranting resection, patients may have concurrent surgical indications that would lead to up-front total thyroidectomy, the nodule may have sonographic features that increase suspicion for malignancy, or surveillance may not be feasible. Some studies have described MT-negative nodules that are resected for clinical reasons, and ideally, this assessment should be made before obtaining MT at all. Consideration of patient preferences, clinical presentation, and cost-effectiveness are all necessary components in shared decision-making.

In summary, the results presented here demonstrate in a hypothetical model that use of either the GSC or TSv3 molecular adjunct for Bethesda III/IV nodules is highly cost effective compared with routine diagnostic thyroid lobectomy for achieving matched therapy and diagnosis. Furthermore, although both MT are associated with cost savings when considering optimally matched therapy and diagnosis, TSv3 was robustly less costly and more likely to yield a correct diagnosis.

### Acknowledgments

We are grateful to James Gallagher, MD, for his critical assessment of the decision model and analysis. This work was supported, in part, by a generous gift from William and Susan Johnson.

### Author Disclosure Statement

The authors have no commercial affiliation associated with this article. Some of the authors (M.S.R., K.L.M., S.E.C., and L.Y.) are employees of the University of Pittsburgh Physicians, which is an affiliate of UPMC. UPMC has granted CBLPath, Inc. a license to market the ThyroSeq trademark for commercial use. The authors receive no direct or indirect compensation related to CBLPath, Inc.

### Funding Information

Research reported in this publication was supported by the National Center for Advancing Translational Sciences of the National Institutes of Health under Award Number

TL1TR001858. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

### References

- Bongiovanni M, Spitale A, Faquin WC, Mazzucchelli L, Baloch ZW 2012 The Bethesda System for Reporting Thyroid Cytopathology: a meta-analysis. *Acta Cytol* **56**:333–339.
- Cooper DS, Doherty GM, Haugen BR, Kloos RT, Lee SL, Mandel SJ, Mazzaferri EL, McIver B, Pacini F, Schlumberger M, Sherman SI, Steward DL, Tuttle RM 2006 Revised American Thyroid Association management guidelines for patients with thyroid nodules and differentiated thyroid cancer. *Thyroid* **19**:1167–1219.
- Cooper DS, Doherty GM, Haugen BR, Kloos RT, Lee SL, Mandel SJ, Mazzaferri EL, McIver B, Pacini F, Schlumberger M, Sherman SI, Steward DL, Tuttle RM 2009 Revised American Thyroid Association management guidelines for patients with thyroid nodules and differentiated thyroid cancer: the American Thyroid Association (ATA) Guidelines Taskforce on Thyroid Nodules and Differentiated Thyroid Cancer. *Thyroid* **19**:1167–1214.
- Sosa JA, Hanna JW, Robinson KA, Lanman RB 2013 Increases in thyroid nodule fine-needle aspirations, operations, and diagnoses of thyroid cancer in the United States. *Surgery* **154**:1420–1426; discussion 1426–1427.
- Nikiforov YE 2017 Role of molecular markers in thyroid nodule management: then and now. *Endocr Pract* **23**:979–988.
- Slotkin JR, Ross OA, Newman ED, Comrey JL, Watson V, Lee RV, Brosious MM, Gerrity G, Davis SM, Paul J, Miller EL, Feinberg DT, Toms SA 2017 Episode-based payment and direct employer purchasing of healthcare services: recent bundled payment innovations and the Geisinger health system experience. *Neurosurgery* **80**:S50–S58.
- Li H, Robinson KA, Anton B, Saldanha IJ, Ladenson PW 2011 Cost-effectiveness of a novel molecular test for cytologically indeterminate thyroid nodules. *J Clin Endocrinol Metab* **96**:E1719–E1726.
- Yip L, Farris C, Kabaker AS, Hodak SP, Nikiforova MN, McCoy KL, Stang MT, Smith KJ, Nikiforov YE, Carty SE 2012 Cost impact of molecular testing for indeterminate thyroid nodule fine-needle aspiration biopsies. *J Clin Endocrinol Metab* **97**:1905–1912.
- Lee L, How J, Tabah RJ, Mitmaker EJ 2014 Cost-effectiveness of molecular testing for thyroid nodules with atypia of undetermined significance cytology. *J Clin Endocrinol Metab* **99**:2674–2682.
- Rivas AM, Nassar A, Zhang J, Casler JD, Chindris AM, Smallridge R, Bernet V 2018 Thyroseq<sup>®</sup> V2.0 molecular testing: a cost-effective approach for the evaluation of indeterminate thyroid nodules. *Endocr Pract* **24**:780–788.
- Labourier E 2016 Utility and cost-effectiveness of molecular testing in thyroid nodules with indeterminate cytology. *Clin Endocrinol (Oxf)* **85**:624–631.
- Wu JX, Lam R, Levin M, Rao J, Sullivan PS, Yeh MW 2016 Effect of malignancy rates on cost-effectiveness of routine gene expression classifier testing for indeterminate thyroid nodules. *Surgery* **159**:118–126.
- Shapiro S, Pharaon M, Kellermeyer B 2017 Cost-effectiveness of gene expression classifier testing of indeterminate thyroid nodules utilizing a real cohort comparator. *Otolaryngol Head Neck Surg* **157**:596–601.
- Balentine CJ, Vanness DJ, Schneider DF 2018 Cost-effectiveness of lobectomy versus genetic testing (Afirma(R))

- for indeterminate thyroid nodules: considering the costs of surveillance. *Surgery* **163**:88–96.
15. Patel KN, Angell TE, Babiarz J, Barth NM, Blevins T, Duh QY, Ghossein RA, Harrell RM, Huang J, Kennedy GC, Kim SY, Kloos RT, LiVolsi VA, Randolph GW, Sadow PM, Shanik MH, Sosa JA, Traweek ST, Walsh PS, Whitney D, Yeh MW, Ladenson PW 2018 Performance of a genomic sequencing classifier for the preoperative diagnosis of cytologically indeterminate thyroid nodules. *JAMA Surg* **153**:817–824.
  16. Nikiforova MN, Mercurio S, Wald AI, Barbi de Moura M, Callenberg K, Santana-Santos L, Gooding WE, Yip L, Ferris RL, Nikiforov YE 2018 Analytical performance of the ThyroSeq v3 genomic classifier for cancer diagnosis in thyroid nodules. *Cancer* **124**:1682–1690.
  17. Steward DL, Carty SE, Sippel RS, Yang SP, Sosa JA, Sips JA, Figge JJ, Mandel S, Haugen BR, Burman KD, Baloch ZW, Lloyd RV, Seethala RR, Gooding WE, Chiosea SI, Gomes-Lima C, Ferris RL, Folek JM, Khawaja RA, Kundra P, Loh KS, Marshall CB, Mayson S, McCoy KL, Nga ME, Ngiam KY, Nikiforova MN, Poehls JL, Ringel MD, Yang H, Yip L, Nikiforov YE 2019 Performance of a multigene genomic classifier in thyroid nodules with indeterminate cytology: a prospective blinded multicenter study. *JAMA Oncol* **5**:204–212.
  18. Haugen BR, Alexander EK, Bible KC, Doherty GM, Mandel SJ, Nikiforov YE, Pacini F, Randolph GW, Sawka AM, Schlumberger M, Schuff KG, Sherman SI, Sosa JA, Steward DL, Tuttle RM, Wartofsky L 2016 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* **26**:1–133.
  19. Lubitz CC, Kong CY, McMahon PM, Daniels GH, Chen Y, Economopoulos KP, Gazelle GS, Weinstein MC 2014 Annual financial impact of well-differentiated thyroid cancer care in the United States. *Cancer* **120**:1345–1352.
  20. Arias E, Heron M, Hu J 2017 United States life tables, 2014. *Natl Vital Stat Rep*. **66**:1–64.
  21. Sanders GD, Neumann PJ, Basu A, Brock DW, Feeny D, Krahn M, Kuntz KM, Meltzer DO, Owens DK, Prosser LA, Salomon JA, Sculpher MJ, Trikalinos TA, Russell LB, Siegel JE, Ganiats TG 2016 Recommendations for conduct, methodological practices, and reporting of cost-effectiveness analyses: second panel on cost-effectiveness in health and medicine. *JAMA* **316**:1093–1103.
  22. Centers for Medicare & Medicaid Services. Medicare reimbursement schedule. Available at [www.cms.gov](http://www.cms.gov) (accessed May 3, 2018).
  23. Nishino M, Nikiforova M 2018 Update on molecular testing for cytologically indeterminate thyroid nodules. *Arch Pathol Lab Med* **142**:446–457.
  24. Alexander EK, Kennedy GC, Baloch ZW, Cibas ES, Chudova D, Diggans J, Friedman L, Kloos RT, LiVolsi VA, Mandel SJ, Raab SS, Rosai J, Steward DL, Walsh PS, Wilde JJ, Zeiger MA, Lanman RB, Haugen BR 2012 Preoperative diagnosis of benign thyroid nodules with indeterminate cytology. *N Engl J Med* **367**:705–715.
  25. Chaudhary S, Hou Y, Shen R, Hooda S, Li Z 2016 Impact of the Afirma gene expression classifier result on the surgical management of thyroid nodules with category III/IV cytology and its correlation with surgical outcome. *Acta Cytol* **60**:205–210.
  26. Samulski TD, LiVolsi VA, Wong LQ, Baloch Z 2016 Usage trends and performance characteristics of a “gene expression classifier” in the management of thyroid nodules: an institutional experience. *Diagn Cytopathol* **44**:867–873.
  27. Baca SC, Wong KS, Strickland KC, Heller HT, Kim MI, Barletta JA, Cibas ES, Krane JF, Marqusee E, Angell TE 2017 Qualifiers of atypia in the cytologic diagnosis of thyroid nodules are associated with different Afirma gene expression classifier results and clinical outcomes. *Cancer Cytopathol* **125**:313–322.
  28. Harrison G, Sosa JA, Jiang X 2017 Evaluation of the Afirma gene expression classifier in repeat indeterminate thyroid nodules. *Arch Pathol Lab Med* **141**:985–989.
  29. Kay-Rivest E, Tibbo J, Bouhabel S, Tamilia M, Leboeuf R, Forest VI, Hier MP, Savoury L, Payne RJ 2017 The first Canadian experience with the Afirma gene expression classifier test. *J Otolaryngol Head Neck Surg* **46**:25.
  30. Valderrabano P, Leon ME, Centeno BA, Otto KJ, Khazai L, McCaffrey JC, Russell JS, McIver B 2016 Institutional prevalence of malignancy of indeterminate thyroid cytology is necessary but insufficient to accurately interpret molecular marker tests. *Eur J Endocrinol* **174**:621–629.
  31. Valderrabano P, Khazai L, Leon ME, Thompson ZJ, Ma Z, Chung CH, Hallanger-Johnson JE, Otto KJ, Rogers KD, Centeno BA, McIver B 2017 Evaluation of ThyroSeq v2 performance in thyroid nodules with indeterminate cytology. *Endocr Relat Cancer* **24**:127–136.
  32. Labourier E, Shifrin A, Busseniers AE, Lupo MA, Manganello ML, Andruss B, Wylie D, Beaudenon-Huibregtse S 2015 Molecular testing for miRNA, mRNA, and DNA on fine-needle aspiration improves the preoperative diagnosis of thyroid nodules with indeterminate cytology. *J Clin Endocrinol Metab* **100**:2743–2750.
  33. Nikiforova MN, Wald AI, Roy S, Durso MB, Nikiforov YE 2013 Targeted next-generation sequencing panel (ThyroSeq) for detection of mutations in thyroid cancer. *J Clin Endocrinol Metab* **98**:E1852–E1860.
  34. Nikiforov YE, Carty SE, Chiosea SI, Coyne C, Duvvuri U, Ferris RL, Gooding WE, LeBeau SO, Ohori NP, Seethala RR, Tublin ME, Yip L, Nikiforova MN 2015 Impact of the multi-gene ThyroSeq next-generation sequencing assay on cancer diagnosis in thyroid nodules with Atypia of undetermined significance/follicular lesion of undetermined significance cytology. *Thyroid* **25**:1217–1223.
  35. Nikiforov YE, Seethala RR, Tallini G, Baloch ZW, Basolo F, Thompson LD, Barletta JA, Wenig BM, Al Ghuzlan A, Kakudo K, Giordano TJ, Alves VA, Khanafshar E, Asa SL, El-Naggar AK, Gooding WE, Hodak SP, Lloyd RV, Maytal G, Mete O, Nikiforova MN, Nose V, Papotti M, Poller DN, Sadow PM, Tischler AS, Tuttle RM, Wall KB, LiVolsi VA, Randolph GW, Ghossein RA 2016 Nomenclature revision for encapsulated follicular variant of papillary thyroid carcinoma: a paradigm shift to reduce overtreatment of indolent tumors. *JAMA Oncol* **2**:1023–1029.

Address correspondence to:

*Linwah Yip, MD*

*Division of Endocrine Surgery*

*Department of Surgery*

*University of Pittsburgh*

*3471 Fifth Avenue Ste 101*

*Pittsburgh, PA 15213*

*E-mail: [yiopl@upmc.edu](mailto:yiopl@upmc.edu)*