Thyroid

DOI: 10.1089/thy.2023.0651

Does the Adoption of Molecular Testing Cause Decreased Thyroidectomy Rates in a National Cohort? A Quasi-Experimental Study of High- Vs. Low-Adoption States

Yongmei Huang, MD*,[1] Stephanie J. Chan, AB*,[2] Jason D. Wright, MD,[1,3] Jennifer H. Kuo, MD,[2] Catherine M. McManus, MD, [2] James A. Lee, MD,[2] Eric J. Kuo, MD[2]

*These two authors contributed equally to this work

[1]Department of Obstetrics and Gynecology, Columbia University, New York, NY

[2]Division of GI/Endocrine Surgery, Columbia University, New York, NY

[3]Herbert Irving Comprehensive Cancer Center, Columbia University, New York, NY

Corresponding author:

Eric J. Kuo, MD

Division of GI and Endocrine Surgery

Columbia University Irving Medical Center

161 Fort Washington Avenue, HIP-8

New York, New York 10032

Telephone: 212-305-0444

Fax: 212-305-0445

Email: ejk2193@cumc.columbia.edu

Twitter: @EricJKuoMD

Thyroid
Does the Adoption of Molecular Testing Cause Decreased Thyroidectomy Rates in a National Cohort? A Quasi-Experimental Study of High- Vs. Low-Adoption States (DOI: 10.1089/thy.2023.0651)
This paper has been peer-reviewed and accepted for publication, but has yet to undergo copyediting and proof correction. The final published version may differ from this proof.

Abstract

Background: Over the last decade, the utilization of molecular testing (MT) for the evaluation of thyroid nodules has increased. Rates and patterns of adoption of MT and its effect on thyroidectomy rates nationally are unknown. Varying rates of MT adoption at the state level provides an opportunity to study the effects of MT on thyroidectomy rates using a quasi-experimental study design.

Methods: We performed a retrospective analysis of American adult patients in the Merative™ MarketScan® Research Databases who underwent thyroid fine needle aspiration (FNA) from 2011-2021. MT included commercially available DNA and RNA platforms and traditional targeted mutational analysis. Interrupted time series analysis was used to evaluate the inflection of MT adoption and thyroidectomy rates after 2015. Difference-in-differences analysis was used to causally analyze the effect of MT adoption on thyroidectomy rates in high- (at least a 10% increase in MT utilization) versus low-adoption states (no more than 5% increase in MT utilization) from 2015 to 2021.

Results: We identified 471,364 patients who underwent thyroid FNA. The utilization of MT increased over the study period from 0.01% (95%CI: 0.00 to 0.02%) to 10.1% (95%CI: 9.7 to 10.5%), in 2021, with an immediate (β2 =1.61, *P*=0.002) and deeper (β3=0.6, *P*<0.001) increase in MT adoption after 2015. Utilization of MT was lower in Black patients, the elderly, rural areas, and patients with Medicaid (*P*<0.05). Thyroidectomy rates were inversely correlated with MT utilization (r=-0.98, *P*<0.0001). From 2015 to 2021, the average MT utilization rate increased from 2.4% to 15.3% in high-adoption states and 1.6% to 5.6% in low-adoption states. In low-adoption states, thyroidectomy rates decreased more but to similar levels (18.5% to 13.2%) compared to high-adoption states (15.9% to 13.4%) with an adjusted difference-in-differences rate of -3.3% (95% CI -5.6% to -0.8%).

Conclusions: The acceleration in adoption of MT after 2015 likely coincides with publication of American Thyroid Association guidelines. Black, elderly, and rural patients are less likely to receive molecular testing. Although thyroidectomy rates were inversely correlated with MT utilization, our study suggests this correlation is not causal. The effect of MT on thyroidectomy rates may be overshadowed by decreasing aggressiveness of thyroid nodule evaluation.

Thyroid
Does the Adoption of Molecular Testing Cause Decreased Thyroidectomy Rates in a National Cohort? A Quasi-Experimental Study of High- Vs. Low-Adoption States (DOI: 10.1089/thy.2023.0651)
This paper has been peer-reviewed and accepted for publication, but has yet to undergo copyediting and proof correction. The final published version may differ from this proof.

Introduction

Thyroid nodules affect over 60% of the population worldwide.[1-2] These common nodules are often asymptomatic and benign: fine needle aspiration (FNA) classifies 55-74% of thyroid nodules as benign while only 2-8% are malignant.[2-3] Notably, 15-30% of thyroid nodules as determined by FNA are cytologically indeterminate.[1,3-4] Indeterminate nodules have historically been referred for diagnostic thyroid surgery.[3,6] However, because indeterminate nodules are often benign, many patients may be exposed to the risks of thyroid surgery with only diagnostic rather than therapeutic benefit.

Molecular testing (MT) has evolved into tool to rule out malignancy in indeterminate thyroid nodules and reduce unnecessary thyroid surgery. A variety of commercial MT platforms based on somatic mutational analysis, RNA/DNA gene expression classifiers, and microRNA evaluation are available, and validation studies show excellent diagnostic performance.[1-3,8-9] Multiple events have influenced the increasing adoption of molecular testing over the last decade, as commercial platforms such as Afirma (Gene Expression Classifier, 2012; Genomic Sequencing Classifier, 2018) and Thyroseq (v1, 2013; v2, 2015; v3 2018) have undergone iterative improvements. Furthermore, in 2015, several significant events with respect to MT in the United States (US) occurred, including the release of Thyroseq v2, the approval of the Afirma Current Procedural Terminology (CPT) code, and incorporation of MT into the American Thyroid Association (ATA) guidelines for thyroid nodules and differentiated thyroid cancer.[7]

While evidence supporting the efficacy of MT for indeterminate thyroid nodules has grown, rates of utilization and factors influencing MT adoption are unknown. Furthermore, while the primary purpose of MT is currently to reduce unnecessary thyroid surgery, it has yet to be demonstrated that MT causes an observable decrease in thyroidectomy rates at a national level. We hypothesized that converging events in 2015 led to heterogeneous state-level MT adoption in the US, and that quasi-experimental methods would be able to demonstrate whether variable MT adoption at the state-level has led to a greater decrease in thyroidectomy rates in high-adoption states vs. low-adoption states.

Thyroid

Does the Adoption of Molecular Testing Cause Decreased Thyroidectomy Rates in a National Cohort? A Quasi-Experimental Study of High- Vs. Low-Adoption States (DOI: 10.1089/thy.2023.0651)

Materials and Methods

Data Source

We performed a retrospective cohort study using the Merative™ MarketScan® Research Databases from 2008 to 2021.[10] The MarketScan® Research Databases contain individual-level, de-identified, healthcare claims information from employers, health plans, hospitals, and Medicare and Medicaid programs (government-funded insurance programs) in the United States. The Commercial Claims and Encounters Database and Medicare Supplemental and Coordination of Benefits databases include over 273 million unique patients since 1995. These data represent the national healthcare experience (both inpatient and outpatient setting) of insured employees and their dependents for active employees, early retirees, and Medicare-eligible retirees with employer-provided Medicare Supplemental plans. The ability to capture patients in both inpatient and outpatient settings is critical to studies that concurrently analyze outpatient evaluation and thyroidectomy, which can be either an inpatient or outpatient procedure. The annual databases include private-sector health data from more than 350 unique payers in all 50 states. The Multi-State Medicaid Database reflects the healthcare service use of beneficiaries covered by Medicaid programs in numerous geographically dispersed states. This study uses de-identified data and was determined to be exempt by the Columbia University Institutional Review Board.

Study Population

We identified patients who underwent thyroid FNA by selecting for patients with CPT and International Classification of Disease (ICD) procedure codes for FNA in combination with diagnosis codes for thyroid nodules or goiter (Table S1). The most recent FNA date was identified as the index FNA date if a patient had multiple FNAs. Patients were eligible to the analytical cohort if they were 18 years or older, had continuous health insurance enrollment from one month before index FNA date to 3 months after.

Thyroid
Does the Adoption of Molecular Testing Cause Decreased Thyroidectomy Rates in a National Cohort? A Quasi-Experimental Study of High- Vs. Low-Adoption States (DOI: 10.1089/thy.2023.0651)
This paper has been peer-reviewed and accepted for publication, but has yet to undergo copyediting and proof correction. The final published version may differ from this proof.

Outcomes and Covariates

Patients were deemed to have undergone MT if they had CPT codes for molecular tests from one month before FNA date to three months after the date of index FNA. Molecular tests included commercially available DNA and RNA platforms in addition to targeted mutational analysis of BRAF, RET/PTC rearrangements, PAX8/PPARG, TERT or other targeted genomic sequence analysis panels (Table S1). Thyroidectomy was identified using ICD and CPT procedure codes, and categorized into three groups: lobectomy, total thyroidectomy, and unspecified thyroidectomy (Table S1). Because cytology results are not available in MarketScan, thyroidectomy rates were defined per FNA rather than per Bethesda III or IV result.

Additional patients' demographics data included health insurance (Medicaid, commercial insurance with/without supplemental Medicare), age at FNA, self-reported gender (male, female), race (White, African-American, Hispanic, other/unknown), metropolitan statistical area (MSA), and insurance type. An MSA is a region with a high population density which is in contrast to rural regions. The information on race was only available in Medicaid dataset; while geographic data, including states where patients resided, was only available for commercial insured patients with/without supplemental Medicare.

Statistical Analyses

The study period was divided into pre- and post-intervention periods coinciding with the publication of 2015 American Thyroid Association Management Guidelines for Adult Patients with Thyroid Nodules and Differentiated Thyroid Cancer on January 12, 2016.[7] Descriptive statistics were used to summarize patients' characteristics depending on their date of FNA (pre-intervention from 2011 to 2015 vs. post-intervention from 2016 to 2021) and MT status (yes vs. no). Standardized mean differences (SMD, the difference in means or proportions divided by the pooled standard error) were used to compare study populations' demographics at pre- and post-2015. We considered a SMD greater than 0.1 to indicate a clinically significant difference between two periods. Chi square tests were used to determine differences in MT rates by patient characteristics. Pearson's correlation

Thyroid

Does the Adoption of Molecular Testing Cause Decreased Thyroidectomy Rates in a National Cohort? A Quasi-Experimental Study of High- Vs. Low-Adoption States (DOI: 10.1089/thy.2023.0651)

coefficient was used to examine a correlation between MT utilization and thyroidectomy rates at both annual and state levels.

Quasi-experimental methods were used to examine the effect of ATA guideline publication on rates of MT and thyroidectomy. Interrupted time series analysis (ITS) evaluates a baseline outcome (β0) and pre-intervention trend (β1) and determines whether the intervention leads to immediate changes in outcome (β2) or changes in post-intervention trends (β3).[11-12] In our study, unadjusted ITS analysis was used to quantify changes in MT and thyroidectomy rates pre-2015 and post-2015. ITS analyses were accomplished using ordinary least squares (OLS) regression with Newey-West autocorrelation adjusted standard errors. A prior defined lag of one year assumed that annual mean molecular testing rate and mean thyroidectomy rate were correlated with the previous one-year rates.

Difference in difference (DID) analysis is a quasi-experimental method that evaluates causality when randomized trials are not feasible. DID analysis defines an intervention effect as a difference between an expected difference in outcomes had an intervention not occurred and actual differences in outcomes after the intervention took place.[13] We applied DID analysis to quantify the effect of MT adoption on thyroidectomy rates in low- vs. high-adoption states. In other words, assuming trends in thyroidectomy rates between states would be constant had MT not been adopted, changes in trends in thyroidectomy rates between states can be causally attributed to variable rates of adoption of MT. As geographic data was not available in the MarketScan Medicaid database, the DID analysis was restricted to commercial insured patients with known information on states where they resided in year 2015 or 2021. The utilization rates of MT and thyroidectomy were calculated at individual state level in 2015 and 2021, respectively. States where MT utilization increased <5% from 2015 to 2021 were defined as low-adoption states while states where MT utilization increased by ≥10% were defined as high-adoption states. Each state had to contribute at least 20 cases of FNA in 2015 and 2021 to be included in the DID analysis. The differences of thyroidectomy rates between 2015 and 2021 and 95% confidence intervals were calculated in low adoption states and high adoption states. A marginal binomial regression model based on generalized estimation

Thyroid
Does the Adoption of Molecular Testing Cause Decreased Thyroidectomy Rates in a National Cohort? A Quasi-Experimental Study of High- Vs. Low-Adoption States (DOI: 10.1089/thy.2023.0651)
This paper has been peer-reviewed and accepted for publication, but has yet to undergo copyediting and proof correction. The final published version may differ from this proof.

equation with the identity linkage was developed at the patient level to calculate the adjusted difference in difference of thyroidectomy rates by including the indicators of pre- or post-period and high or low adoption states and interaction between these two indicators as well as adjusting for patients' age, gender, MSA, and region. If our hypothesis were true in terms of increasing MT adoption causing decreasing thyroidectomy rates, we would observe a greater decrease in thyroidectomy rate in high-adoption states compared to low-adoption states, with the point estimate as a positive delta (difference) in the difference of post- and pre- thyroidectomy rates with the 95% confidence intervals not crossing zero.

All hypothesis tests were two-sided. A P-value of <0.05 was considered statistically significant. All analyses were conducted using SAS version 9.4 (SAS Institute Inc, Cary, North Carolina). The study adhered to the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) guidelines.

Results

We identified 807,877 patients who underwent thyroid FNA from 2008-2021 (Figure 1). After excluding patients with incomplete enrollment, patients with age < 18 years, and patients from 2008-2010 where the MT rate was 0%, our final cohort included 471,364 patients who underwent thyroid FNA from 2011 to 2021. Of these, 66,515 (14.1%) underwent multiple FNAs. 83.7% of patients were female and the median age at FNA was 52 years (Table 1).

Overall, 16,218 (3.6%) patients underwent MT and 71,640 (15.2%) underwent thyroidectomy. From 2011 to 2021, the utilization of MT as a proportion of patients undergoing thyroid FNA increased from 0.01% (95%CI: 0.00-0.02%) to 10.1% (95%CI: 9.7 to 10.5%), while the thyroidectomy rate decreased from 17.5% (95%CI: 17.2 to 17.8%) to 12.5% (95%CI: 12.1 to 13.0%) (Figure 2A). Afirma claims were most common (N=8964), followed by TMA (N=7721), Thyroseq (N=1427), and ThyGeNext/ThyraMIR N=612 (Figure 2B). Thyroidectomy rates by MT sub-type were comparable, although an increase in thyroidectomy rate with TMA compared to the commercially available MTs was observed from 2018-2021 (Figure 2C). Adoption of the various subtypes of MT were similar across

regions, with the exception of Thyroseq which was predominantly used in the Northeast (Figure 3).

A negative correlation between MT and thyroidectomy rates were observed at both the annual level (r=-0.98, *P* < 0.001) for the overall study population and at the state level (r=-0.28, *P*=0.03) for commercially-insured patients (Figure 4). In examining characteristics associated with MT, we found that patients who received MT were more frequently younger (median age 50 [IQR 41-58] years vs. 52 [IQR 42-59] years), male (4.2% vs. 3.3% female), Hispanic (4.7% vs 2.2% White vs. 1.4% Black), had commercial insurance (3.5% vs 2.1% Medicaid), and resided in a MSA (3.5% vs 2.4% non-MSA) (Table 2).

85.1% (13,808 of 16,218) of the MTs in our study took place after 2015, coinciding with publication of ATA guidelines (Table 1). Interrupted time series analysis demonstrated an immediate increase in MT utilization after 2015 (β2=1.61, SE 0.32, *P*=0.002) (Table 3, Figure 5A). Furthermore, the increasing trend of MT utilization rate post-2015 exceeded that of the pre-2015 era (β3=0.60, SE 0.09, P<0.001). In contrast, thyroidectomy rates did not immediately change post-2015 (β2=-0.50, SE 0.30, P=0.138) and continued to decrease at a similar slope as the pre-2015 trend (β3=-0.025, SE 0.128, P=0.851) (Table 3, Figure 5B).

We identified 17 high-adopting states (N=16,273) where MT utilization increased by at least 10% and 8 low-adopting states (N=5,756) where MT utilization increased by no more than 5% (Table 4). From 2015 to 2021, the average MT utilization rate increased from 2.4% to 15.3% in high-adoption states and 1.6% to 5.6% in low-adoption states (Figure 6). In contrast to our hypothesis that thyroidectomy rates would decrease more in high-adoption states, we observed that the decrease in thyroidectomy rates in low-adoption states (-5.3%, 95% CI -3.5 to -1.9) exceeded that of high-adoption states (-2.5%, 95% CI -3.6 to -1.3). The unadjusted difference-in-difference was of -2.6% (95% CI -3.5 to -1.9). After controlling for age, sex, MSA, and region, we observed an adjusted difference-in-differences of -3.2% (95% CI -0.8 to -5.6). However, 2021 thyroidectomy rates in both high- (13.4%) and low- (13.2%) adoption states were comparable.

Thyroid
Does the Adoption of Molecular Testing Cause Decreased Thyroidectomy Rates in a National Cohort? A Quasi-Experimental Study of High- Vs. Low-Adoption States (DOI: 10.1089/thy.2023.0651)
This paper has been peer-reviewed and accepted for publication, but has yet to undergo copyediting and proof correction. The final published version may differ from this proof.

Discussion

This retrospective national cohort study outlined trends in MT adoption over the last decade. In addition to identifying several populations where MT is less frequently utilized, we demonstrated a steady increase in MT utilization and decrease in thyroidectomy rates since 2011, with a specific inflection point denoting increased MT adoption around 2015. However, in contrast to our hypothesis, we observed greater decreases in thyroidectomy rates in low-adoption states compared to high-adoption states, suggesting a correlative rather than causal relationship between MT utilization and thyroidectomy.

Literature surrounding MT has largely focused on performance and cost-effectiveness, while data on its utilization and adoption are lacking. We believe our study is the first to show the rate at which MT utilization has increased in the United States, from 0% to approximately 10% nationally over the last decade. However, given that rates of indeterminate cytology are even higher, ranging from 15-30%, our study shows that MT technology has yet to fully saturate clinical practice on a national level.[1,3-4,7,14-15]

There are several possible reasons why MT utilization has not fully saturated clinical practice in the US. Our study suggests that factors often associated with limited access to health care also hinder access to MT. We observed that MT is less frequently utilized by under-insured, elderly, rural, and Black patients. Commercial MT platforms cost in excess of $3000 per test and coverage by insurance can be variable. Although Hispanic patients more frequently received molecular testing, race data was only available for Medicaid patients, and therefore race observations may not be generalizable to the entire population. Disparities have been similarly demonstrated across other aspects of thyroid cancer care in the United States, from diagnosis to treatment. Minority race and lower socioeconomic status have been associated with advanced thyroid cancer stage at presentation.[16-18] Rates of postoperative complications after thyroidectomy are higher in racial/ethnic minority groups,[16,19] and patients with lower educational attainment are more likely to encounter inadequate thyroid cancer care.[16,20] MT therefore represents an

additional area, with unique challenges due to its financial cost and variability in insurance coverage, where conscious effort must be paid to decreasing disparities in care.

The lack of full MT adoption may also be explained by skepticism regarding the practical impact of MT on surgical decision-making. Several single-institution investigations have challenged the clinical utility of molecular profiling for indeterminate thyroid nodules.[15,21-22] A 2016 study from Noureldine et al. found that MT altered surgical decision making in only 7.9% of patients who underwent MT; in other words, other factors such as imaging characteristics and patient preference were concordant with MT results such that the added value of MT was small.[15] Specific situations in which MT was found to be redundant included highly suspicious sonographic features suggesting malignancy, large nodules causing compressive symptoms, and patient preference for surgery and against surveillance. Similarly, Huang et al. determined that MT, in relation to its preceding diagnostic steps (i.e. patient interview, sonographic evaluation, and FNA biopsy), did not significantly improve upon the ability to distinguish malignant thyroid nodules.[21] Another study of a high-volume thyroid center noted unchanged thyroidectomy rates even after adoption of MT, challenging the utility of MT.[22] Therefore, while the performance of MT has been validated and the negative predictive value sufficiently high to rule out malignancy, there may be a discordance between the actual and perceived performance of MT such that a significant amount of surgical decision-making nationally may be occurring independently of MT results.

Furthermore, it is important to understand the utility of MT in the context of its cost. Cost effectiveness analyses (CEAs) calculate an incremental cost-effectiveness ratio (ICER), which is the quotient of incremental changes in costs over incremental changes in outcome, typically quality-adjusted life years (QALYs). An assumption of CEAs analyzing MT is that given an intervention of MT or no MT, all differences in outcome are attributed to MT when in fact, MT may only alter decision making in a fraction of these cases, as demonstrated by Noureldine et al.[23-28] In other words, the benefit in outcome attributed to MT may be overestimated, and when the incremental benefit provided by MT is considered, the ICER may become less favorable for MT. In a comparison of reflexive versus selective strategies of MT, selective MT was the more cost-effective strategy if the

Thyroid
Does the Adoption of Molecular Testing Cause Decreased Thyroidectomy Rates in a National Cohort? A Quasi-Experimental Study of High- Vs. Low-Adoption States (DOI: 10.1089/thy.2023.0651)
This paper has been peer-reviewed and accepted for publication, but has yet to undergo copyediting and proof correction. The final published version may differ from this proof.

costs of MT exceeded $1,050, a threshold which current costs of MTs in the US do exceed.[29]

Despite these obstacles, our study shows that MT rates have increased nationally in the US after the publication of the first large, multicenter validation study of MT in 2012.[8] Although adoption subsequently increased, our study suggests an additional inflection point occurred in 2015, likely representing the composite effect of the 2015 ATA guidelines, approval of the Afirma CPT code, and release of Thyroseq v2 on clinical practice.[7] Previous studies have also documented the impact of the 2015 ATA guidelines on practice patterns. For instance, implementation of the more conservative 2015 ATA guideline-concordant treatment for low-risk patients with well-differentiated thyroid carcinomas resulted in decreased rates of up-front total thyroidectomies and completion thyroidectomies.[14,31] Use of radioactive iodine therapy after total thyroidectomies also decreased following publication of the guidelines.[14] Notably, although our study demonstrated an immediate and sustained increase in MT utilization, thyroidectomy rates remained in stable decline without a change in the trend post-2015. This stable decline may be attributable to a general decrease in aggressiveness in the US approach to the evaluation of thyroid nodules, as evidenced by increasing sonographic thresholds for fine needle aspiration, lobectomy versus total thyroidectomy, and active surveillance for papillary microcarcinoma.

Discordance between the post-2015 acceleration of MT adoption and lack of equivalent post-2015 deceleration of thyroidectomy rates was also demonstrated in our difference-in-differences analysis. We paradoxically demonstrated that low-adoption states had a larger decrease in thyroidectomy rates compared to high-adoption states. This suggests that the correlation we observed between declining thyroidectomy rates and increasing MT utilization may be correlative rather than causal and could be explained by other factors. However, absolute thyroidectomy rates in low- and high-adoption states in 2021 were comparable, suggesting that an overall decrease in aggressiveness towards pursuing thyroid surgery perhaps overshadowed this effect between 2015 and 2021.

Does the Adoption of Molecular Testing Cause Decreased Thyroidectomy Rates in a National Cohort? A Quasi-Experimental Study of High- Vs. Low-Adoption States (DOI: 10.1089/thy.2023.0651)
This paper has been peer-reviewed and accepted for publication, but has yet to undergo copyediting and proof correction. The final published version may differ from this proof.

Several limitations to this study require acknowledgment. Administrative datasets are generally limited by missing data, coding errors, and inadequate granularity. Particularly relevant is lack of cytology results, requiring our analysis to evaluate MT and thyroidectomy rates as a proportion of FNAs as a whole rather than the subset with indeterminate cytology. However, given that rates of indeterminate cytology may be increasing, even approaching nearly 50% at an academic medical center, the number of total FNAs may be a more stable denominator with which to normalize MT utilization rates.[30] MT utilization was derived from claims data, and because CPT codes for the different MT subtypes may vary by payer and the CPT codes available at the time, there may be mixing and migration between MT subtypes within our cohort, which is why these subtypes were analyzed collectively rather than individually. Marketscan is largely limited to commercial insurers, which limits the applicability of our results across the range of insurers in the US. Due to heterogeneity of MT adoption within states, our state-level analysis perhaps obscured effects that would have possibly been observed on an institutional-level. However, codes to identify individual treating facilities are not captured in Marketscan, which likely diluted the capture of a real effect of MT on thyroidectomy rates.

In summary, there has been an increase in MT utilization and decrease in thyroidectomies over the past decade nationwide. The MT utilization inflection point at 2015 likely exemplifies how the recent ATA guidelines have directly influenced clinical practice patterns. The growing adoption of MT leads us to confront the true efficacy and cost-effectiveness of MT for betterment of patient care and the healthcare system more broadly. As we determine the conditions in which MT is best utilized, there also must be awareness and action against disparities in access to MT. Future studies of the real-world cost of MT, as well as analysis of the effects of MT on an institutional-level rather than a state level, may improve our understanding of MT's true impact on thyroid nodule evaluation.

Thyroid
Does the Adoption of Molecular Testing Cause Decreased Thyroidectomy Rates in a National Cohort? A Quasi-Experimental Study of High- Vs. Low-Adoption States (DOI: 10.1089/thy.2023.0651)
This paper has been peer-reviewed and accepted for publication, but has yet to undergo copyediting and proof correction. The final published version may differ from this proof.

Acknowledgments

This work was presented as a poster at the American Thyroid Association Annual Meeting 2023 in Washington, DC.

Authorship Confirmation/Contribution Statement

Yongmei Huang: Formal Analysis, Methodology, Writing-Review & Editing; **Stephanie J. Chan**: Writing-Original Draft; **Jason D. Wright**: Writing-Review & Editing; **Jennifer H. Kuo**: Conceptualization, Writing-Review & Editing; **Catherine M. McManus**: Writing-Review & Editing; **James A. Lee**: Writing-Review & Editing; **Eric J. Kuo**: Conceptualization, Methodology, Writing-Original Draft, Writing-Review & Editing

Disclosures

The authors have no financial conflicts of interest to disclose.

Funding/Financial Support

This research did not receive any specific grant from funding agencies in the public, commercial, or non-profit sectors.

References

1. Roth MY, Witt RL, Steward DL. Molecular testing for thyroid nodules: Review and current state. Cancer 2018;124(5):888–898; doi: 10.1002/cncr.30708.
2. Grani G, Sponziello M, Pecce V, et al. Contemporary Thyroid Nodule Evaluation and Management. The Journal of Clinical Endocrinology & Metabolism 2020;105(9):2869–2883; doi: 10.1210/clinem/dgaa322.
3. Alexander EK, Cibas ES. Diagnosis of thyroid nodules. The Lancet Diabetes & Endocrinology 2022;10(7):533–539; doi: 10.1016/S2213-8587(22)00101-2.
4. Bose S, Sacks W, Walts AE. Update on Molecular Testing for Cytologically Indeterminate Thyroid Nodules. Advances in Anatomic Pathology 2019;26(2):114–123; doi: 10.1097/PAP.0000000000000211.
5. Cibas ES, Ali SZ. The 2017 Bethesda System for Reporting Thyroid Cytopathology. Journal of the American Society of Cytopathology 2017;6(6):217–222; doi: 10.1016/j.jasc.2017.09.002.
6. Sauter JL, Lehrke H, Zhang X, et al. Assessment of The Bethesda System for Reporting Thyroid Cytopathology: Surgical and Long-Term Clinical Follow-up of 2,893 Thyroid Fine-Needle Aspirations. American Journal of Clinical Pathology 2019;152(4):502–511; doi: 10.1093/ajcp/aqz076.
7. Haugen BR, Erik K Alexander, Bible, Keith C., 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; doi: 10.1089/thy.2015.0020.
8. Alexander EK, Kennedy GC, Baloch ZW, et al. Preoperative Diagnosis of Benign Thyroid Nodules with Indeterminate Cytology. New England Journal of Medicine 2012;367(8):705–715; doi: 10.1056/NEJMoa1203208.
9. Livhits MJ, Zhu CY, Kuo EJ, et al. Effectiveness of Molecular Testing Techniques for Diagnosis of Indeterminate Thyroid Nodules: A Randomized Clinical Trial. JAMA Oncology 2021 ;7(1) :70–77 ; doi : 10.1001/jamaoncol.2020.5935.
10. MarketScan Research Databases – Databases | IBM. In.

Thyroid

Does the Adoption of Molecular Testing Cause Decreased Thyroidectomy Rates in a National Cohort? A Quasi-Experimental Study of High- Vs. Low-Adoption States (DOI: 10.1089/thy.2023.0651)

This paper has been peer-reviewed and accepted for publication, but has yet to undergo copyediting and proof correction. The final published version may differ from this proof.

11. Bernal JL, Cummins S, Gasparrini A. Interrupted time series regression for the evaluation of public health interventions: a tutorial. Int J Epidemiol 2017;46(1):348–355; doi: 10.1093/ije/dyw098.
12. Penfold RB, Zhang F. Use of Interrupted Time Series Analysis in Evaluating Health Care Quality Improvements. Academic Pediatrics 2013;13(6, Supplement):S38–S44; doi: 10.1016/j.acap.2013.08.002.
13. Wing C, Simon K, Bello-Gomez RA. Designing Difference in Difference Studies: Best Practices for Public Health Policy Research. Annu Rev Public Health 2018;39:453–469; doi: 10.1146/annurev-publhealth-040617-013507.
14. Gordon AJ, Dublin JC, Patel E, et al. American Thyroid Association Guidelines and National Trends in Management of Papillary Thyroid Carcinoma. JAMA Otolaryngology–Head & Neck Surgery 2022;148(12):1156–1163; doi: 10.1001/jamaoto.2022.3360.
15. Noureldine SI, Najafian A, Aragon Han P, et al. Evaluation of the Effect of Diagnostic Molecular Testing on the Surgical Decision-Making Process for Patients With Thyroid Nodules. JAMA Otolaryngology–Head & Neck Surgery 2016;142(7):676–682; doi: 10.1001/jamaoto.2016.0850.
16. Chen DW, Lang BHH, McLeod DSA, et al. Thyroid cancer. The Lancet 2023;401(10387):1531–1544; doi: 10.1016/S0140-6736(23)00020-X.
17. Weeks KS, Kahl AR, Lynch CF, et al. Racial/ethnic differences in thyroid cancer incidence in the United States, 2007-2014. Cancer 2018;124(7):1483–1491; doi: 10.1002/cncr.31229.
18. Nnorom SO, Baig H, Akinyemi OA, et al. Persistence of Disparity in Thyroid Cancer Survival After Adjustments for Socioeconomic Status and Access. The American Surgeon™ 2022;88(7):1484–1489; doi: 10.1177/00031348221082282.
19. Kovatch KJ, Reyes-Gastelum D, Hughes DT, et al. Assessment of Voice Outcomes Following Surgery for Thyroid Cancer. JAMA Otolaryngology–Head & Neck Surgery 2019;145(9):823–829; doi: 10.1001/jamaoto.2019.1737.
20. Bilimoria KY, Bentrem DJ, Linn JG, et al. Utilization of total thyroidectomy for papillary thyroid cancer in the United States. Surgery 2007;142(6):906-913.e2; doi: 10.1016/j.surg.2007.09.002.

Thyroid

Does the Adoption of Molecular Testing Cause Decreased Thyroidectomy Rates in a National Cohort? A Quasi-Experimental Study of High- Vs. Low-Adoption States (DOI: 10.1089/thy.2023.0651)

This paper has been peer-reviewed and accepted for publication, but has yet to undergo copyediting and proof correction. The final published version may differ from this proof.

21. Huang BL, Chabot JA, Lee JA, et al. A stepwise analysis of the diagnostic algorithm for the prediction of malignancy in thyroid nodules. Surgery 2020;167(1):28–33; doi: 10.1016/j.surg.2019.05.079.
22. Sacks WL, Bose S, Zumsteg ZS, et al. Impact of Afirma gene expression classifier on cytopathology diagnosis and rate of thyroidectomy. Cancer Cytopathology 2016;124(10):722–728; doi: 10.1002/cncy.21749.
23. Dharampal N, Smith K, Harvey A, et al. Cost-effectiveness analysis of molecular testing for cytologically indeterminate thyroid nodules. J Otolaryngol Head Neck Surg 2022;51(1):46; doi: 10.1186/s40463-022-00604-7.
24. Labourier E. Utility and cost-effectiveness of molecular testing in thyroid nodules with indeterminate cytology. Clin Endocrinol (Oxf) 2016;85(4):624–631; doi: 10.1111/cen.13096.
25. Lee L, How J, Tabah RJ, et al. Cost-effectiveness of molecular testing for thyroid nodules with atypia of undetermined significance cytology. J Clin Endocrinol Metab 2014;99(8):2674–2682; doi: 10.1210/jc.2014-1219.
26. Nicholson KJ, Roberts MS, McCoy KL, et al. Molecular Testing Versus Diagnostic Lobectomy in Bethesda III/IV Thyroid Nodules: A Cost-Effectiveness Analysis. Thyroid 2019;29(9):1237–1243; doi: 10.1089/thy.2018.0779.
27. 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; doi: 10.1177/0194599817725709.
28. Rivas AM, Nassar A, Zhang J, et al. ThyroSeq®V2.0 Molecular Testing: A Cost-Effective Approach for the Evaluation of Indeterminate Thyroid Nodules. Endocr Pract 2018;24(9):780–788; doi: 10.4158/EP-2018-0212.
29. Hu QL, Schumm MA, Zanocco KA, et al. Cost analysis of reflexive versus selective molecular testing for indeterminate thyroid nodules. Surgery 2022;171(1):147–154; doi: 10.1016/j.surg.2021.04.050.
30. Ramonell KM, Ohori NP, Liu JB, et al. Changes in thyroid nodule cytology rates after institutional implementation of the Thyroid Imaging Reporting and Data System. Surgery 2023;173(1):232–238; doi: 10.1016/j.surg.2022.06.061.

31. Hirshoren N, Kaganov K, Weinberger JM, et al. Thyroidectomy Practice After Implementation of the 2015 American Thyroid Association Guidelines on Surgical Options for Patients With Well-Differentiated Thyroid Carcinoma. JAMA Otolaryngology–Head & Neck Surgery 2018;144(5):427–432; doi: 10.1001/jamaoto.2018.0042.

Thyroid
Does the Adoption of Molecular Testing Cause Decreased Thyroidectomy Rates in a National Cohort? A Quasi-Experimental Study of High- Vs. Low-Adoption States (DOI: 10.1089/thy.2023.0651)
This paper has been peer-reviewed and accepted for publication, but has yet to undergo copyediting and proof correction. The final published version may differ from this proof.

Table 1. Patients' demographics by pre- and post- intervention (guidance released)

	Overall		Pre-intervention (2011-2015)		Post-intervention (2016-2021)		SMD
	N	%	N	%	N	%	
Total patients	471364	100	279714	59.3	191650	40.7	
Health insurance							
Medicaid	55965	11.9	21650	7.7	34315	17.9	-0.3076
Commercial	415399	88.1	258064	92.3	157335	82.1	
Age at FNA							0.0317
Median (IQR)	52	42-59	52	42-59	52	42-59	
18-29	26911	5.7	16017	5.7	10894	5.7	
30-39	66521	14.1	39271	14	27250	14.2	
40-49	111522	23.7	66873	23.9	44649	23.3	
50-59	159400	33.8	96033	34.3	63367	33.1	
60-64	83069	17.6	48559	17.4	34510	18	
65-91	16849	3.6	10055	3.6	6794	3.6	
Sex							-0.0018
Male	76737	16.3	44936	16.1	31801	16.6	
Female	394627	83.7	234778	83.9	159849	83.4	
Race							
White	28720	51.3	11129	51.4	17591	51.3	0.0771
Black	17028	30.4	6900	31.9	10128	29.5	
Hispanic	1579	2.8	455	2.1	1124	3.3	
Other/Unknown	8638	15.4	3166	14.6	5472	16.0	

Thyroid
Does the Adoption of Molecular Testing Cause Decreased Thyroidectomy Rates in a National Cohort? A Quasi-Experimental Study of High- Vs. Low-Adoption States (DOI: 10.1089/thy.2023.0651)
This paper has been peer-reviewed and accepted for publication, but has yet to undergo copyediting and proof correction. The final published version may differ from this proof.

Thyroid
Does the Adoption of Molecular Testing Cause Decreased Thyroidectomy Rates in a National Cohort? A Quasi-Experimental Study of High- Vs. Low-Adoption States (DOI: 10.1089/thy.2023.0651)

Metropolitan statistics area							
Non-MSA	48348	11.6	32191	12.5	16157	10.3	0.4540
MSA	341429	82.2	222640	86.3	118789	75.5	
Unknown	25622	6.2	3233	1.3	22389	14.2	
Region							
Northeast	100253	24.1	65900	25.5	34353	21.8	0.2307
North Central	76725	18.5	48417	18.8	28308	18.0	
South	169069	40.7	96568	37.4	72501	46.1	
West	65533	15.8	43747	17.0	21786	13.9	
Unknown	3819	0.9	3432	1.3	387	0.3	
Molecular testing							
No	455146	96.4	277304	99.1	177842	92.8	
Yes	16218	3.6	2410	0.9	13808	7.2	
Surgery							
None	399724	84.8	233830	83.6	165894	86.6	
Lobectomy	28350	6.0	28336	10.1	14910	7.8	
Total thyroidectomy	43246	9.2	17545	6.3	10805	5.6	
Unspecific thyroidectomy	44	0.0	3	0.0	41	0.0	

MSA, metropolitan statistical area; *FNA*, fine needle aspiration; *SMD,* standardized mean difference

Notes: The information on race was only available in Medicaid dataset, and geographic data was only available for commercial insured patients and Medicare beneficiaries. Standardized mean differences (SMD) greater than 0.1 indicate a clinically significant difference between pre- and post-intervention groups.

Table 2. Patients' demographics by molecular testing

	Molecular Testing						*p*-value
	No			Yes			
	N	Col %	Row %	N	Col %	Row %	
Total patients	455146	100.0	96.6	16218	100.0	3.4	
Health insurance							<0.0001
Medicaid	54790	12.0	97.9	1175	7.2	2.1	
Commercial	400356	88.0	96.4	15043	92.8	3.6	
Age at FNA							<0.0001
Median (IQR)	52		(42-59)	50		(41-58)	
18-29	25862	5.7	96.1	1049	6.5	3.9	
30-39	63838	14	96.0	2683	16.5	4.0	
40-49	107556	23.6	96.4	3966	24.5	3.6	
50-59	153944	33.8	96.6	5456	33.6	3.4	
60-64	80343	17.7	96.7	2726	16.8	3.3	
65-91	16561	3.6	98.3	288	1.8	1.7	
Sex							<0.0001
Male	73482	16.1	95.8	3255	20.1	4.2	
Female	381664	83.9	96.7	12963	79.9	3.3	
MSA							<0.0001
Missing	54790	12	97.9	1175	7.2	2.1	
Non-MSA	47210	10.4	97.7	1138	7	2.4	
MSA	329529	72.4	96.5	11900	73.4	3.5	
Unknown	23617	5.2	92.2	2005	12.4	7.8	

Thyroid
Does the Adoption of Molecular Testing Cause Decreased Thyroidectomy Rates in a National Cohort? A Quasi-Experimental Study of High- Vs. Low-Adoption States (DOI: 10.1089/thy.2023.0651)
This paper has been peer-reviewed and accepted for publication, but has yet to undergo copyediting and proof correction. The final published version may differ from this proof.

Region							<0.0001
Missing	54790	12	97.9	1175	7.2	2.1	
Northeast	95314	20.9	95.1	4939	30.5	4.9	
North Central	74753	16.4	97.4	1972	12.2	2.6	
South	162863	35.8	96.3	6206	38.3	3.7	
West	63657	14	97.1	1876	11.6	2.9	
Unknown	3769	0.8	98.7	50	0.3	1.3	
Race							<0.0001
Missing	400356	88	96.4	15043	92.8	3.6	
White	28089	6.2	97.8	631	3.9	2.2	
Black	16797	3.7	98.6	231	1.4	1.4	
Hispanic	1505	0.3	95.3	74	0.5	4.7	
Other/Unknown	8399	1.8	97.2	239	1.5	2.8	
Molecular testing							
Afirma	NA	NA	NA	8964	55.27	NA	
Thyroseq	NA	NA	NA	1427	8.8	NA	
ThyGeNext/ThyraMIR	NA	NA	NA	612	3.8	NA	
TMA	NA	NA	NA	7721	47.6	NA	
Surgery							
None	388616	85.4	NA	11108	68.5	NA	
Lobectomy	26202	5.8	NA	2148	13.2	NA	
Total thyroidectomy	40286	8.9	NA	2960	18.3	NA	
Unspecific thyroidectomy	42	<0.1	NA	2	<0.1	NA	

MSA, metropolitan statistical area; *FNA*, fine needle aspiration; *TMA,* targeted mutational analysis

Thyroid
Does the Adoption of Molecular Testing Cause Decreased Thyroidectomy Rates in a National Cohort? A Quasi-Experimental Study of High- Vs. Low-Adoption States (DOI: 10.1089/thy.2023.0651)
This paper has been peer-reviewed and accepted for publication, but has yet to undergo copyediting and proof correction. The final published version may differ from this proof.

Thyroid

Does the Adoption of Molecular Testing Cause Decreased Thyroidectomy Rates in a National Cohort? A Quasi-Experimental Study of High- Vs. Low-Adoption States (DOI: 10.1089/thy.2023.0651)

Table 3. Interrupted time series analysis examining the association of the guidance release with utilization of molecular testing and thyroidectomy at state level

Variable	Coefficient (SE)	*P* Value
Mean molecular testing rate		
Intercept β0	-0.62(0.14)	0.003
Baseline trend β1	0.53 (0.04)	<0.001
Level change after schedule change β2	1.61 (0.32)	0.002
Trend change after schedule change β3	0.60 (0.09)	<0.001
Mean thyroidectomy rate		
Intercept β0	17.67 (0.29)	<0.001
Baseline trend β1	-0.45 (0.09)	0.002
Level change after schedule change β2	-0.50 (0.30)	0.138
Trend change after schedule change β3	-0.025 (0.128)	0.851

SE, standard error

Notes: Interrupted time series analyses were accomplished using ordinary least squares (OLS) regression with Newey-West autocorrelation adjusted standard errors. A prior defined lag of one year indicated that annual mean molecular testing rate and mean thyroidectomy rate were correlated with the rates one year previous.

Thyroid

Does the Adoption of Molecular Testing Cause Decreased Thyroidectomy Rates in a National Cohort? A Quasi-Experimental Study of High- Vs. Low-Adoption States (DOI: 10.1089/thy.2023.0651)

Table 4. Comparisons of patients' demographics within high adoption and low adoption states by pre- and post-interruption

	High Adoption States (N=17) No. of patients (N=16,273)			Low Adoption States (N=8) No. of patients (N=5,756)		
	Pre-interruption (2015) N=11,113 N (%)	Post-interruption (2020) N=5,160 N (%)	SMD	Pre-interruption (2015) N=4,421 N (%)	Post-interruption (2020) N=1,335 N (%)	SMD
Age at FNA			0.025			-0.031
Median (IQR)	51 (42, 58)	51 (42, 59)		52 (43, 58)	51 (42, 58)	
18-29	619 (5.6)	250 (4.8)		260 (5.9)	83 (6.2)	
30-39	1559 (14.0)	736 (14.3)		582 (13.2)	170 (12.7)	
40-49	2814 (25.3)	1268 (24.6)		1020 (23.1)	352 (26.4)	
50-59	3893 (35.0)	1825 (35.4)		1597 (36.1)	468 (35.1)	
60-64	1955 (17.6)	994 (19.3)		885 (20.0)	225 (16.9)	
65-	273 (2.5)	87 (1.7)		77 (1.7)	37 (2.8)	
Sex			-0.032			0.014
Male	1896 (17.1)	954 (18.5)		705 (16.0)	251 (18.8)	
Female	9217 (82.9)	4206 (81.5)		3716 (84.1)	1084 (81.2)	
Metropolitan statistics area			0.420			0.594
Non-MSA	1010 (9.1)	357 (6.9)		806 (18.2)	209 (15.7)	
MSA	10103 (90.9)	4387 (85.0)		3615 (81.8)	921 (69.0)	
Unknown	0 (0.0)	416 (8.1)		0 (0.0)	205 (15.4)	

Region			0.167		0.361
Northeast	3164 (28.5)	1461 (28.3)	514 (11.6)	187 (14.0)	
North Central	1633 (14.7)	796 (15.4)	413 (9.3)	277 (20.8)	
South	3806 (34.3)	2088 (40.5)	3494 (79.0)	871 (65.2)	
West	2510 (22.6)	815 (15.8)	0 (0.0)	0 (0.0)	

MSA, metropolitan statistical area; *FNA*, fine needle aspiration; *IQR*, interquartile range

Notes: Standardized mean differences (SMD) greater than 0.1 indicate a clinically significant difference between pre- and post-intervention groups.

Thyroid
Does the Adoption of Molecular Testing Cause Decreased Thyroidectomy Rates in a National Cohort? A Quasi-Experimental Study of High- Vs. Low-Adoption States (DOI: 10.1089/thy.2023.0651)
This paper has been peer-reviewed and accepted for publication, but has yet to undergo copyediting and proof correction. The final published version may differ from this proof.

Thyroid
Does the Adoption of Molecular Testing Cause Decreased Thyroidectomy Rates in a National Cohort? A Quasi-Experimental Study of High- Vs. Low-Adoption States (DOI: 10.1089/thy.2023.0651)
This paper has been peer-reviewed and accepted for publication, but has yet to undergo copyediting and proof correction. The final published version may differ from this proof.

Figure Legends

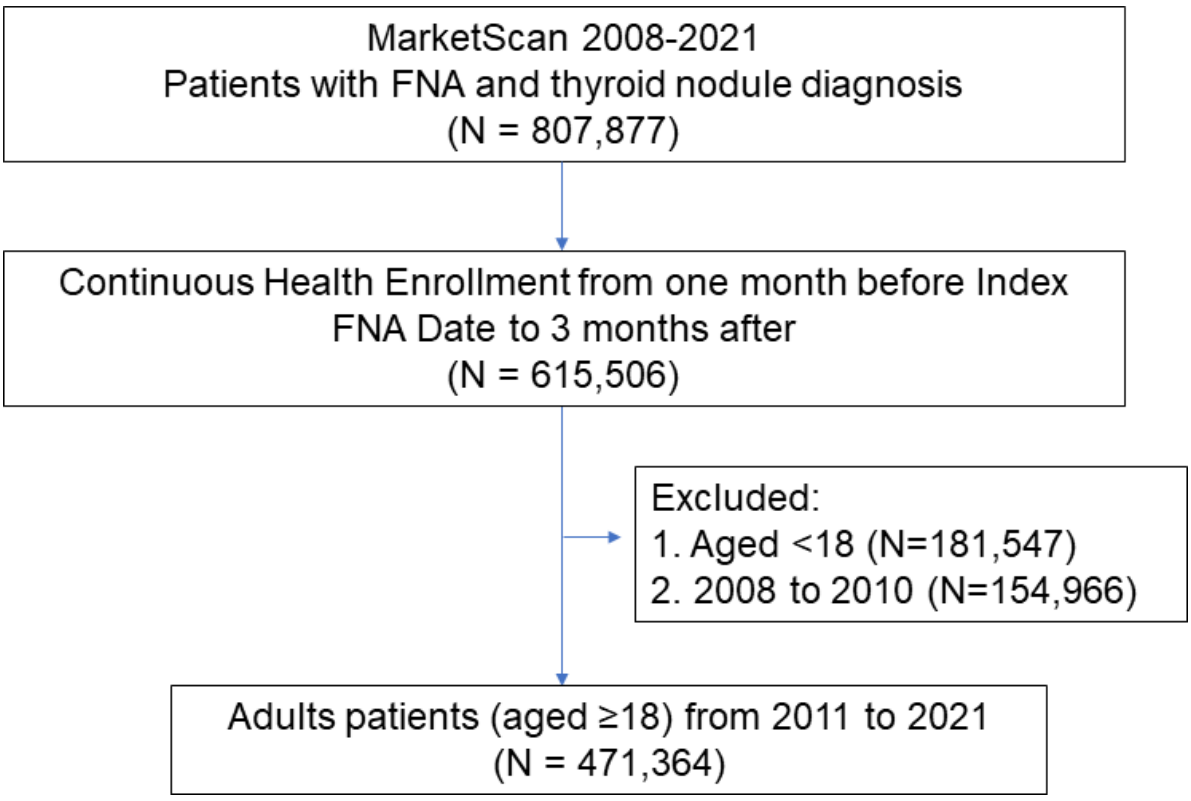

Figure 1. Cohort selection.

FNA, fine needle aspiration

Thyroid

Does the Adoption of Molecular Testing Cause Decreased Thyroidectomy Rates in a National Cohort? A Quasi-Experimental Study of High- Vs. Low-Adoption States (DOI: 10.1089/thy.2023.0651)

This paper has been peer-reviewed and accepted for publication, but has yet to undergo copyediting and proof correction. The final published version may differ from this proof.

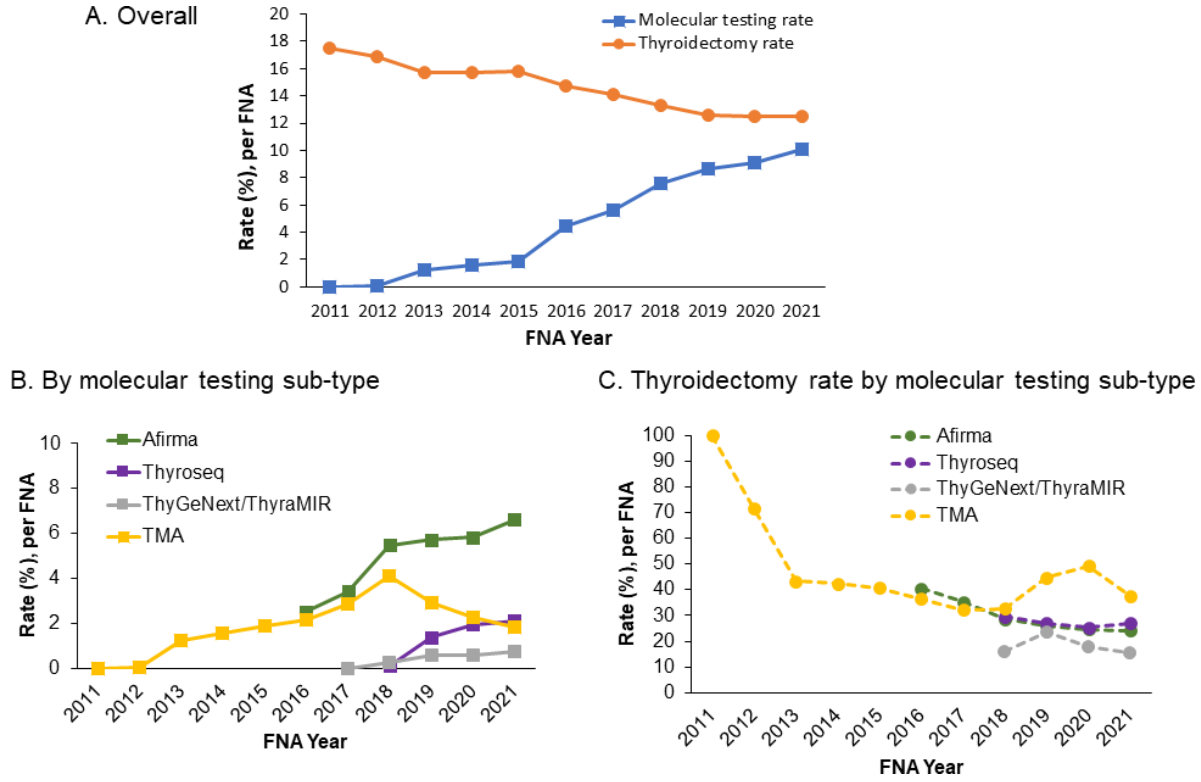

Figure 2. Annual rates of molecular testing utilization and thyroidectomy.

A. Overall rates; B. By molecular testing sub-type; C. Thyroidectomy rates by molecular test subtype (Afirma, N=8964; Thyroseq, N=1427; ThyGeNext/ThyraMIR, N=612; TMA, N=7721)

FNA, fine needle aspiration; TMA, targeted mutational analysis

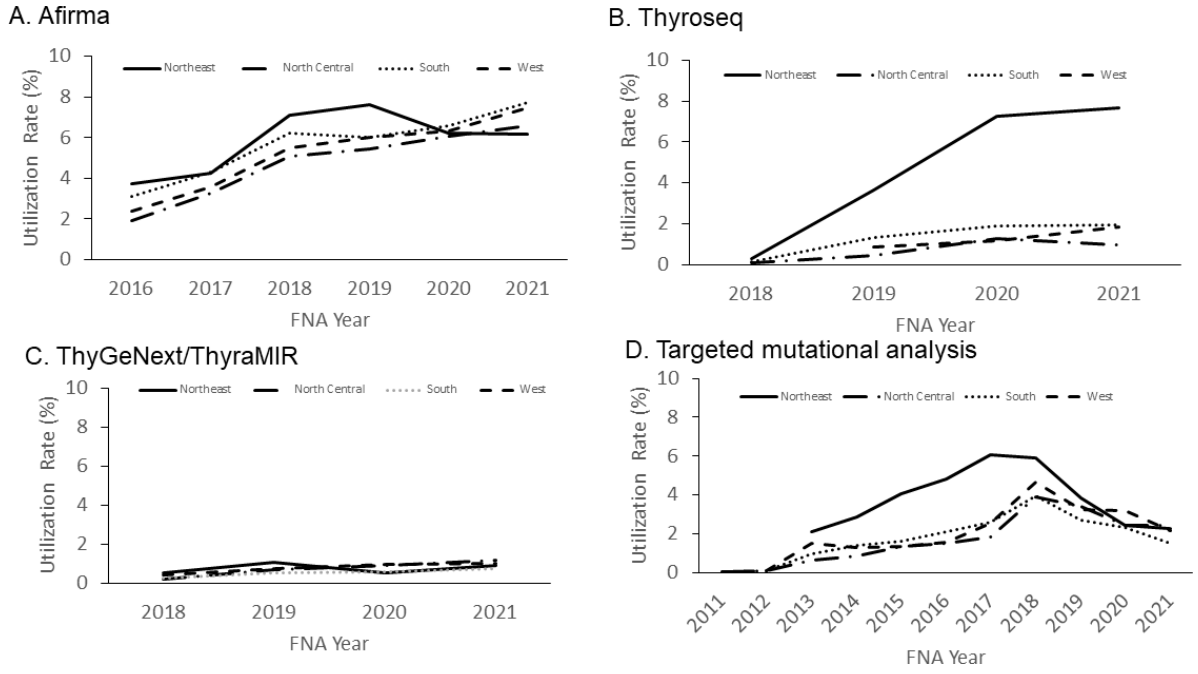

Figure 3. Adoption of molecular tests by region

A. Afirma; B. Thyroseq; C. ThyGeNext/ThyraMIR; D. Targeted mutational analysis

Thyroid
Does the Adoption of Molecular Testing Cause Decreased Thyroidectomy Rates in a National Cohort? A Quasi-Experimental Study of High- Vs. Low-Adoption States (DOI: 10.1089/thy.2023.0651)
This paper has been peer-reviewed and accepted for publication, but has yet to undergo copyediting and proof correction. The final published version may differ from this proof.

Thyroid
Does the Adoption of Molecular Testing Cause Decreased Thyroidectomy Rates in a National Cohort? A Quasi-Experimental Study of High- Vs. Low-Adoption States (DOI: 10.1089/thy.2023.0651)

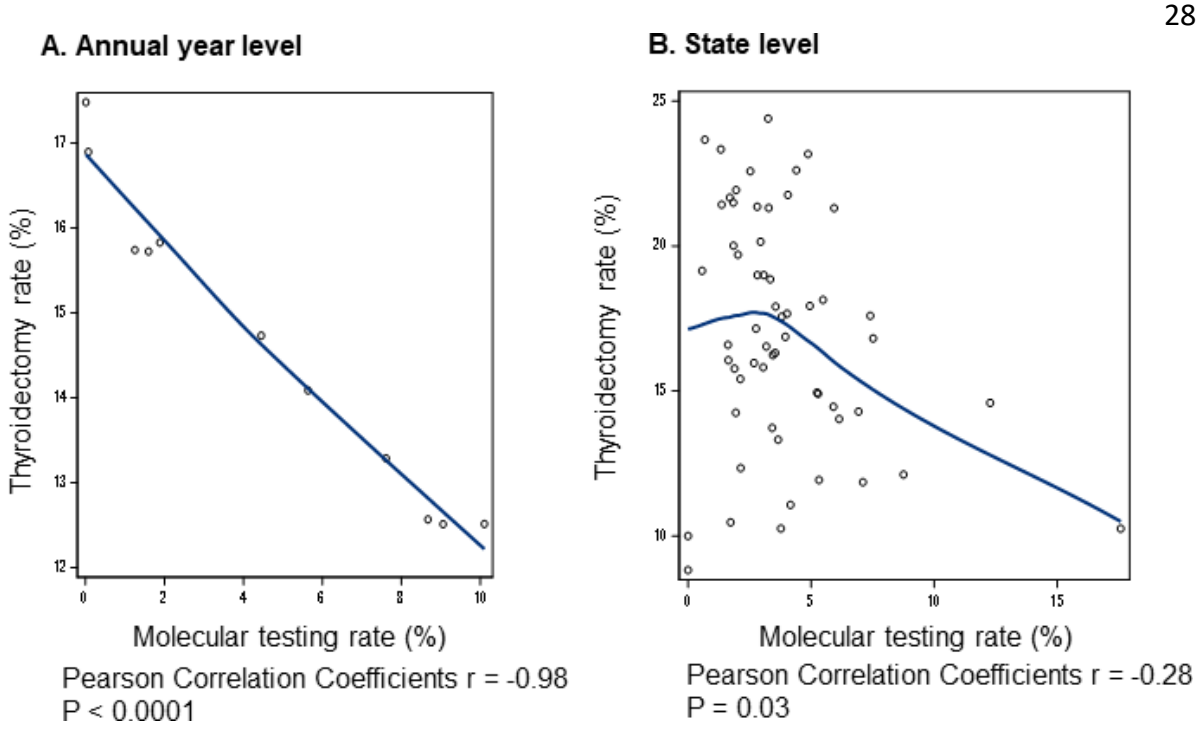

Figure 4. Pearson correlation of molecular testing utilization with thyroidectomy rates at annual year level (A), state level (B).

Thyroid
Does the Adoption of Molecular Testing Cause Decreased Thyroidectomy Rates in a National Cohort? A Quasi-Experimental Study of High- Vs. Low-Adoption States (DOI: 10.1089/thy.2023.0651)
This paper has been peer-reviewed and accepted for publication, but has yet to undergo copyediting and proof correction. The final published version may differ from this proof.

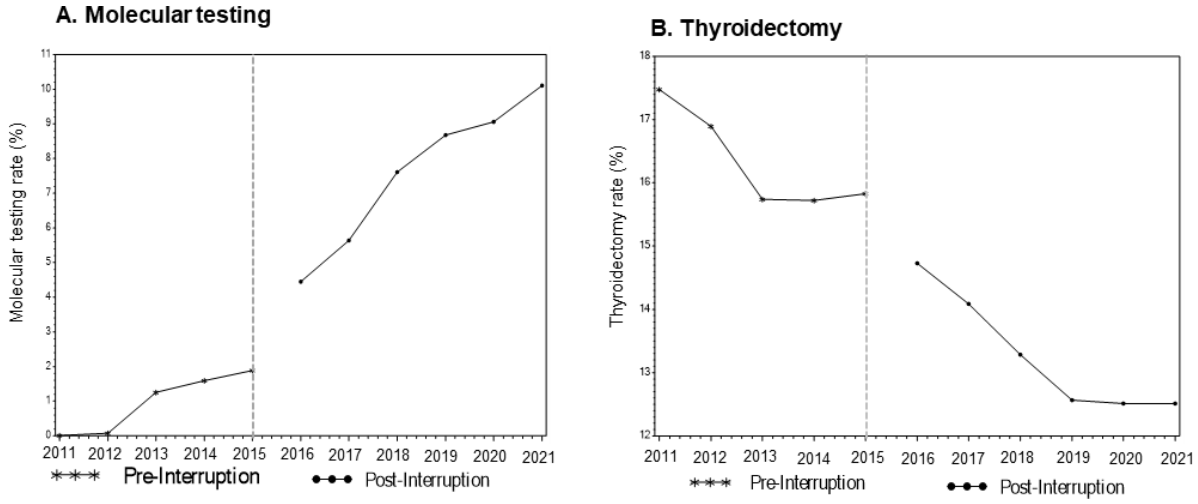

Figure 5. Single interrupted time series analysis for molecular testing (A) and thyroidectomy (B).

Thyroid
Does the Adoption of Molecular Testing Cause Decreased Thyroidectomy Rates in a National Cohort? A Quasi-Experimental Study of High- Vs. Low-Adoption States (DOI: 10.1089/thy.2023.0651)
This paper has been peer-reviewed and accepted for publication, but has yet to undergo copyediting and proof correction. The final published version may differ from this proof.

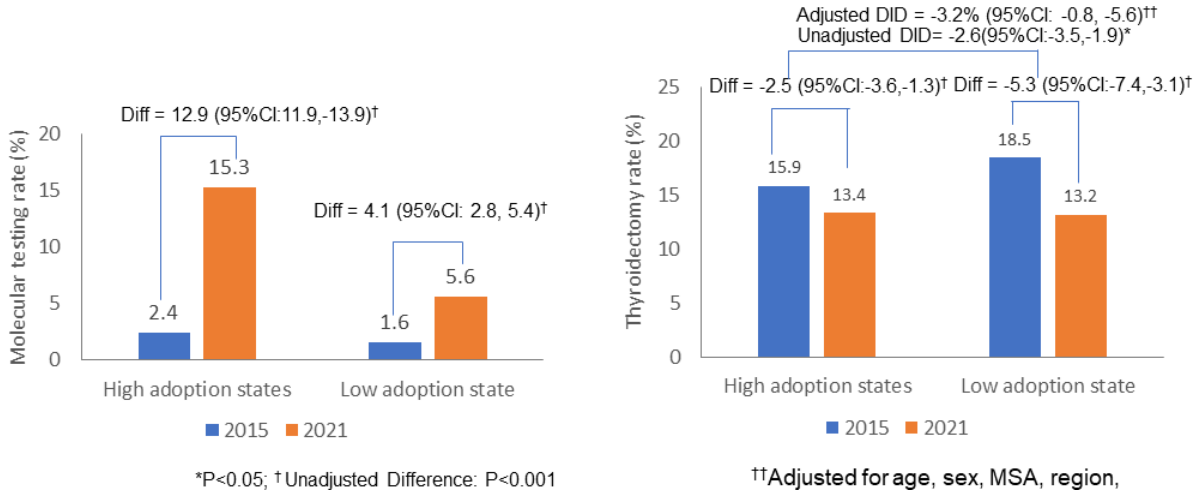

Figure 6. Molecular testing and surgery rates in high-utilizers vs low-utilizers from 2015 to 2021 (N=22,029, high-adoption state N=16,273, low-adoption state N=5,756).

MSA, metropolitan statistical area; DID, difference-in-differences