

# Correlation Between Histological Diagnosis and Mutational Panel Testing of Thyroid Nodules: A Two-Year Institutional Experience

Rupendra T. Shrestha,<sup>1</sup> Maria R. Evasovich,<sup>2</sup> Khalid Amin,<sup>3</sup> Angela Radulescu,<sup>1</sup> Tina S. Sanghvi,<sup>4</sup> Andrew C. Nelson,<sup>3</sup> Maryam Shahi,<sup>3</sup> and Lynn A. Burmeister<sup>1</sup>

**Background:** Indeterminate thyroid fine-needle aspiration (FNA) cytology, including atypia of undetermined significance (AUS/FLUS) and suspicious for follicular neoplasm (SFN), continues to generate uncertainty about the presence of malignancy, resulting in repeated follow-up, repeat FNA, or diagnostic surgery. Mutational panel testing may improve the malignancy risk prediction in indeterminate nodules, but the general application of such testing has not been investigated extensively.

**Methods:** A retrospective review was performed of all patients undergoing thyroidectomy at a tertiary care facility over a two-year period. Mutational panel test results, when present, were analyzed relative to FNA cytologic result and surgical histopathologic diagnosis. Malignancy rates, sensitivity, specificity, positive predictive values (PPV), negative predictive values (NPV) and positive and negative likelihood ratios (LR) were calculated.

**Results:** A total of 261 operated thyroid nodules had the following initial FNA cytology results: 2% non-diagnostic, 23% benign, 28% AUS/FLUS, 11% SFN, 9% suspicious for malignancy (SUSP), and 27% malignant. The histopathologic malignancy rate was 48%, subcategorized by cytology into benign 7%, AUS/FLUS 30%, SFN 38%, and SUSP 83%. Mutations were more frequent in indeterminate nodules that were histologically malignant versus benign ( $p < 0.0001$ ) or versus adenoma ( $p = 0.001$ ). Mutational analysis in 44 AUS/FLUS nodules resulted in a malignancy detection sensitivity of 85%, a specificity of 65%, a PPV of 50%, a NPV of 91%, and a positive LR of 2.4. In 12 SFN nodules analyzed with ThyroSeq<sup>®</sup> testing, sensitivity was 100%, specificity 57%, PPV 63%, NPV 100%, and LR 2.3. Performance of the seven-gene mutational panel was not significantly different from the ThyroSeq<sup>®</sup> panel in the AUS/FLUS group. The malignancy yield, comparing the mutation positive AUS/FLUS group with the untested AUS/FLUS surgical cohort, did not reach statistical significance ( $p = 0.17$ ).

**Conclusions:** In a surgical cohort, a similar NPV but a lower PPV was found with the use of mutational panel testing compared to the published literature. Following the identification of a mutation, the prevalence of malignancy in the AUS/FLUS or SFN category was increased by nearly 15% to 45% and 53%, respectively. Further study is needed to confirm these results and to analyze clinical outcome subcategories relative to the utility of mutational testing.

## Introduction

CYTOLOGICAL EXAMINATION OF fine-needle aspiration (FNA) biopsy samples has been used to increase the diagnostic yield of malignancy detected at surgery, reducing unnecessary surgeries since the 1970s (1,2). Under the currently recommended Bethesda System for Reporting Thyroid Cytopathology (BSRTC) (3), the implied malignancy risk

varies across the continuum of six categories, ranging from 3% with benign cytologic diagnosis to 99% with malignant cytologic diagnosis. However, an undesirably high rate of uncertainty remains in the indeterminate “gray zone” categories, especially with atypia of undetermined significance/follicular lesion of undetermined significance (AUS/FLUS) and suspicious for follicular neoplasm (SFN). To address these uncertainties, management recommendations from the

Departments of <sup>1</sup>Medicine, Division of Diabetes, Endocrinology and Metabolism; <sup>2</sup>Surgery; <sup>3</sup>Laboratory Medicine and Pathology; and <sup>4</sup>Radiology; University of Minnesota, Minneapolis, Minnesota.

BSRTC (3) and recent guidelines have included repeat FNA for nodules with AUS/FLUS cytology and diagnostic lobectomy for SFN cytology (4–6). Advances in understanding of the molecular pathophysiology of thyroid cancer (7–9) have been accompanied by the development of molecular assays, which may aid in establishing a preoperative cancer diagnosis when thyroid cytology specimens yield cytologically indeterminate results (6,10).

Thyroid cancer directed multigene panels for the detection of specific mutations from FNA cytology specimens have been reported by a few groups (11–18). Next-generation sequencing (NGS) has revolutionized DNA sequencing technology, making it cheaper and faster (19). The ThyroSeq<sup>®</sup> panel is based on the NGS analysis of mutations and gene fusions known to occur in >90% of thyroid cancers, and has been reported to have a high negative predictive value (NPV) as well as a high positive predictive value (PPV) for cancer detection in thyroid nodules with AUS/FLUS and SFN (BSRTC III and IV) cytology (15–17). While the reported data on molecular panel testing appear promising in predicting the risk of thyroid cancer (20), the published NGS ThyroSeq<sup>®</sup> literature so far is from a single group (15–17). The clinical utility, when and how to use such testing in indeterminate (specifically AUS/FLUS and FN) categories, has not been well established (6,10,21,22). This study reports on the authors' institutional experience with molecular testing during the first two years of its implementation.

## Materials and Methods

Following Institutional Review Board approval, the records were retrospectively reviewed of all patients undergoing thyroidectomy over a two-year period (January 2013–December 2014) at the University of Minnesota Medical Center (UMMC), a tertiary care University teaching hospital. Baseline demographics, initial FNA cytology results, surgical histopathology, and mutational analysis, when present, were collected. During much of 2013–2014, a dedicated sample was collected and saved for potential future molecular testing on every FNA performed. If the FNA cytology result was read by the pathologist as AUS/FLUS, SFN, or suspicious for malignancy (SUSP), the dedicated sample was automatically sent for mutational analysis as part of an institutional care pathway and, once resulted, was available for the pathologist to review. Due to the high risk of malignancy in the SUSP category, the reflex automatic molecular testing in this group was soon discontinued from the care pathway. A small number of samples were not sent for molecular testing either due to inadequate sampling, or during a short time subinterval where an order was required by the provider. Also, some of the cytology FNAs would have been obtained prior to the start date of the surgical review, or at outside institutions, but were sent to UMMC for second review. These FNAs formed a group that did not undergo molecular testing and were used for comparison to the group that underwent mutational testing.

The mutational analysis was performed at the University of Pittsburgh Medical Center, Division of Molecular and Genomic Pathology. From January of 2013 to September of 2013, seven-gene mutational testing (MT) was performed that utilized real-time LightCycler polymerase chain reaction and fluorescence melting curve analysis to detect possible

mutations (*BRAF*<sup>V600E</sup>, *NRAS* codon 61, *HRAS* codon 61, *KRAS* codons 12 and 13) (14) and single-step real-time reverse transcription real-time polymerase chain reaction to amplify fusion points of the rearrangements (*RET/PTC1*, *RET/PTC3*, and *PAX8/PPARG*). Subsequently, ThyroSeq<sup>®</sup> versions 1 and 2, using NGS platforms, were used as they were made available from the reference laboratory (15). ThyroSeq<sup>®</sup> v1 included NGS for 284 mutations in 12 key thyroid cancer related genes including *AKT1*, *CTNNB1*, *GNAS*, *NRAS*, *HRAS*, *KRAS*, *PIK3CA*, *PTEN*, *RET*, *TP53*, and *TSHR*, and detection of chromosomal rearrangements *RET/PTC1*, *RET/PTC3*, and *PAX8/PPARG*. ThyroSeq<sup>®</sup> v2 included the above point mutations plus *EIF1AX*, *TERT*, and 42 gene fusions involving the following genes: *RET*, *PPARG*, *NTRK1*, *NTRK3*, *ALK*, *BRAF*, and *IGF2BP3*.

Results were analyzed individually based on the initial seven-gene MT or ThyroSeq<sup>®</sup> version and combining the methods. Since the two ThyroSeq<sup>®</sup> versions are similar, except that additional mutations and fusions are available in ThyroSeq<sup>®</sup> v2, and the fact that we detected only a few mutations with ThyroSeq<sup>®</sup> v2, the results of ThyroSeq are reported together. Five samples lacked enough DNA for meaningful interpretation and were classified in the group that did not obtain mutational testing. Each FNA cytology result was correlated to the surgical pathology outcome of the aspirated nodule only. Sub-centimeter nodules that were found to be malignant on surgical histopathology were categorized as malignant. However, if the aspirated nodule was benign but the patient had an incidental cancer, it was categorized as benign for this analysis. Individual molecular mutations and fusions were noted. A chi-square test was performed to determine differences in group frequency. A *p*-value of <0.05 was considered significant. Sensitivity, specificity, PPV, and NPV were calculated for the whole sample and for ThyroSeq<sup>®</sup>. The positive and negative likelihood ratio (LR) was also calculated to estimate the risk of malignancy for any individual with a positive or negative test in the AUS or SFN category.

## Results

Between January 2013 and December 2014, a total of 299 thyroid surgeries were performed. Thirty-eight surgeries performed for Graves' disease, large goiters with or without compressive symptoms, or for completion thyroidectomy following an initial partial thyroidectomy were excluded. The remaining 261 surgeries performed for thyroid nodules were further analyzed for this study.

The demographic data (Table 1) for the patients undergoing surgery were typical of thyroid nodule and cancer populations (5,6). The majority of the patients were women. While 86% of the surgeries were performed by a single surgeon, multiple pathologists interpreted the FNA and histopathologic results.

At the final histopathology, 135/261 (52%) of the operated nodules were nonmalignant (50 nodules were read as adenomas, while 85 had other benign diagnoses), whereas 126/261 (48%) were malignant (Table 1). The majority of malignancies were papillary thyroid carcinomas (including variants; 108/126; 86%), followed by follicular thyroid carcinoma (12/126; 9.5%), medullary thyroid carcinoma (5/126; 4%), and lymphoma (1/126; 0.8%).

TABLE 1. DEMOGRAPHIC DATA OF ALL STUDY SUBJECTS UNDERGOING THYROID SURGERY

Age, <i>M</i> ( <i>SD</i> )	48 (15)
Sex ratio, F:M	3.4: 1
BMI, <i>M</i> ( <i>SD</i> )	29.3 (7.3)
<i>Final histopathological diagnosis of study population</i> (n = 261)	
Benign	135 (52%)
Malignant	126 (48%)
<i>Classification of malignant histopathology</i> (n = 126)	
Papillary thyroid carcinoma	108 (86%)
Follicular thyroid carcinoma	12 (9.5%)
Medullary thyroid carcinoma	5 (4%)
Lymphoma	1 (1%)
<i>Classification of benign histopathology</i> (n = 135)	
Adenoma	50 (37%)
Multinodular goiter/nodular hyperplasia	65 (48%)
Other (benign thyroid tissue, Hashimoto, adenomatoid, etc.)	20 (15%)

Data based on 261 subjects who underwent surgery. Papillary thyroid carcinoma includes both classical and variants.

BMI, body mass index; *SD*, standard deviation.

The cytologic and histologic diagnoses are tabulated and correlated in Table 2. Of the 261 surgeries, 125 (48%) were performed following indeterminate (AUS/FLUS, SFN, and SUSP) cytologic diagnosis, 70 (27%) followed malignant cytologic diagnoses, 60 (23%) followed benign cytologic diagnoses, and 6 (2%) followed non-diagnostic cytologic diagnoses (Table 2). All nodules with malignant FNA on cytology were malignant on final surgical histopathology. Of the 60 nodules with benign FNA cytology, 4 (7%) were malignant at surgery. Nodule growth and large size were the major reasons for surgery in the benign FNA group. In the AUS/FLUS and FN category, 22/73 (30%) and 11/29 (38%) were malignant, respectively, while 19/23 (83%) with SUSP cytology were malignant. None of six non-diagnostic nodules were malignant at surgery.

Mutational testing was performed on 73/125 (58%) nodules in the indeterminate cytology category of which five were non-diagnostic due to lack of sufficient DNA sample (Table 3). Of the remaining 68 nodules, 44 mutational tests were performed for AUS/FLUS (Fig. 1), 12 for SFN (Fig. 2), and 12 for SUSP FNA (Fig. 3) cytologic diagnoses. In most cases, the decision for or against molecular testing was due to

the time period of the study, since indeterminate AUS/FLUS and SFN samples were automatically sent for molecular testing for a period of time covered in the analysis. The seven-gene MT was utilized in 23 samples, and Thyroseq<sup>®</sup> was performed in 45 samples (Table 3).

The overall malignancy rate at surgery was 27/68 (40%) in the indeterminate group that had molecular testing, and 25/57 (44%) in the group that did not undergo molecular testing ( $p=0.64$ ). Among the 68 tested with a mutational panel, a mutation in one or more genes was identified in 24/27 histologically malignant nodules (89%) compared with 8/18 (44%) nodules diagnosed as adenomas and 6/23 (26%) other benign lesions. All mutation-positive histologically benign nodules were described as nodular hyperplasia. The positive mutation rate was higher in the malignant versus benign group ( $p<0.0001$ ) and in the malignant versus follicular adenoma group ( $p=0.001$ ), indicating that mutations are significantly more frequent in malignant nodules.

Within the AUS/FLUS category alone (Fig. 1), 44 nodules had mutational panel testing performed, while the remaining 29 were not tested. The malignancy rate in the tested group was 13/44 (30%), while the untested group had a malignancy rate of 9/29 (31%;  $p=0.89$ ). The mutational panel showed a false-positive result in 11 cases, false negative in two cases, true positive in 11 cases, and true negative in 20 cases. With these results, the sensitivity of the test was calculated to be 85%, specificity 65%, PPV 50%, NPV 91%, positive LR 2.4, and negative LR 0.2. The seven-gene MT was performed in 17 nodules with seven positive mutations (one *BRAF*, three *NRAS*, one *PAX/PPARG* fusion, and two *HRAS*) of which four were malignant, three were adenomas, and one was benign. In this group, three were true positives, four false positives, one false negative, and nine true negatives. ThyroSeq<sup>®</sup> was performed in the remaining 27 nodules, with 15 testing positive for mutations (three *BRAF*, including one that also had *PI3CA*, *TERT*, and *AKT1* mutation in the same nodule; six *NRAS*, including one with a *TSHR* mutation; two *HRAS*; one *TP53*; one *KRAS*; one *EIF1AX*; and two *RET*, of which one also had a *HRAS* mutation). For those undergoing analysis with ThyroSeq<sup>®</sup>, eight were true positive, seven were false positives, one was false negative, and 11 were true negative. The sensitivity and specificity rates did not vary significantly when the two methods were compared. *RAS* mutations were the most common, followed by *BRAF*. While all *BRAF* mutation-positive nodules were histologically

TABLE 2. CYTOLOGIC AND HISTOLOGIC CORRELATES OF THE SURGICAL POPULATION

<i>Cytologic diagnosis</i>	<i>N</i> (% of total operations)	<i>Malignant histology</i> (% of cytologic category)	<i>Adenoma histology</i> (% of cytologic category)
Non-diagnostic	6 (2%)	Non-malignant	1 (17%)
Benign	60 (23%)	4 (7%)	12 (20%)
<b>AUS/FLUS</b>	73 (28%)	22 (30%)	25 (34%)
<b>SFN</b>	29 (11%)	11 (38%)	11 (38%)
<b>SUSP</b>	23 (9%)	19 (83%)	1 (4%)
Malignant	70 (27%)	70 (100%)	0
Total	261	126 (48%)	50 (19%)

Distribution of malignancy and adenoma across Bethesda System for Reporting Thyroid Cytopathology categories. Bold represents indeterminate cytologic categories.

AUS/FLUS, atypia of undetermined significance/follicular lesion of undetermined significance; SFN, suspicious for follicular neoplasm; SUSP, suspicious for malignancy.

TABLE 3. INDETERMINATE CYTOLOGY GROUP ANALYSIS BY MUTATIONAL TESTING

n	Final histological diagnosis			
	125	Malignant	Adenoma	Benign
No mutational testing, <i>n</i> (% of total indeterminate)	52	25 (44%)	19 (33%)	8 (14%)
Mutation DNA insufficient, <i>n</i>	5	0	0	5
Mutational testing, <i>n</i> (% of total indeterminate)	68	27 (41%) <sup>a</sup>	18 (26%)	23 (33%)
Mutation positive, <i>n</i> (% of mutation tested histological category)		24/27 (89%) <sup>b,c</sup>	8/18 (44%) <sup>d</sup>	6/23 (26%)
Seven-gene mutational test	23	6	7	10
ThyroSeq <sup>®</sup>	45	17	11	17

Breakdown of the molecular testing method, if used, in the 125 indeterminate cytology nodules as presented in area in bold in Table 2.

<sup>a</sup>*p*=0.74 vs. the malignant group that did not get molecular testing.

<sup>b</sup>*p*<0.0001 malignant vs. (adenoma + benign).

<sup>c</sup>*p*=0.001 vs. adenoma in the molecular tested groups.

<sup>d</sup>*p*=0.21 adenoma vs. benign in the molecular tested group.

malignant, only a third of the nine *NRAS* mutations were found in malignant lesions, two out of four *HRAS* mutations were detected in malignant nodules, and one nodule with a *KRAS* mutation was read as benign on final histopathologic diagnosis.

Within the SFN category (Fig. 2), 12/29 nodules were tested with ThyroSeq<sup>®</sup> while 17/29 did not undergo molecular testing. The malignancy rate in the tested group was 5/12 (42%), while the untested group had a malignancy rate of 6/17 (35%). There was no significant difference in the malignancy rates in these two groups (*p*=0.73). Of the 12 tested lesions, five were true positive, three were false positive, none had a false-negative result, and four were true negatives. With these results, the sensitivity was 100%, specificity 57%, PPV 63%, NPV 100%, positive LR 2.3, and negative LR 0. All mutations were found in *RAS* (*NRAS*, *KRAS*, or *HRAS*) genes.

For the 23 SUSP nodules (Fig. 3), 12 underwent mutational testing, while 11 did not. Nine of 12 (75%) were found to be malignant in the group that underwent testing, while 10/11

(91%) were ultimately malignant in the group that did not undergo testing. This difference was not statistically significant (*p*=0.31). There were eight true positives, no false positives, one false negative, and three true negatives. All but one mutation in the SUSP category were in the *BRAF* gene, the other being a *RAS* mutation.

In this series, *RAS* mutations (*NRAS*, *HRAS*, and *KRAS*) were the most commonly detected alteration (total 23) followed by mutations in *BRAF*. All 11 nodules that had *BRAF* mutations were malignant. The overall malignancy rate for *NRAS*-positive nodules was 6/14 (43%), for *HRAS* 2/5 (40%), and for *KRAS* 3/4 (75%). Other mutations were uncommon. A single sub-centimeter nodule with aggressive biological behavior carried four mutations (*BRAF*, *TERT*, *PIK3CA*, and *AKT1*) (23). Of the two *RET* mutations identified, one was seen in a medullary thyroid carcinoma, whereas the other was associated with a *HRAS* mutation and histopathologically diagnosed follicular adenoma. Another nodule with a *TSHR* mutation combined with an *NRAS* mutation was histopathologically diagnosed as a follicular

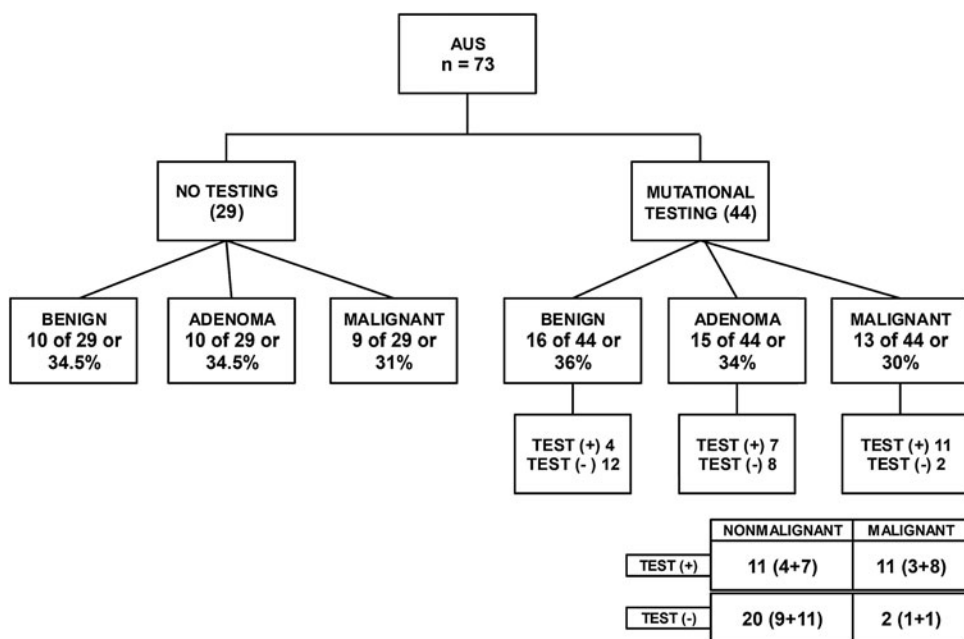
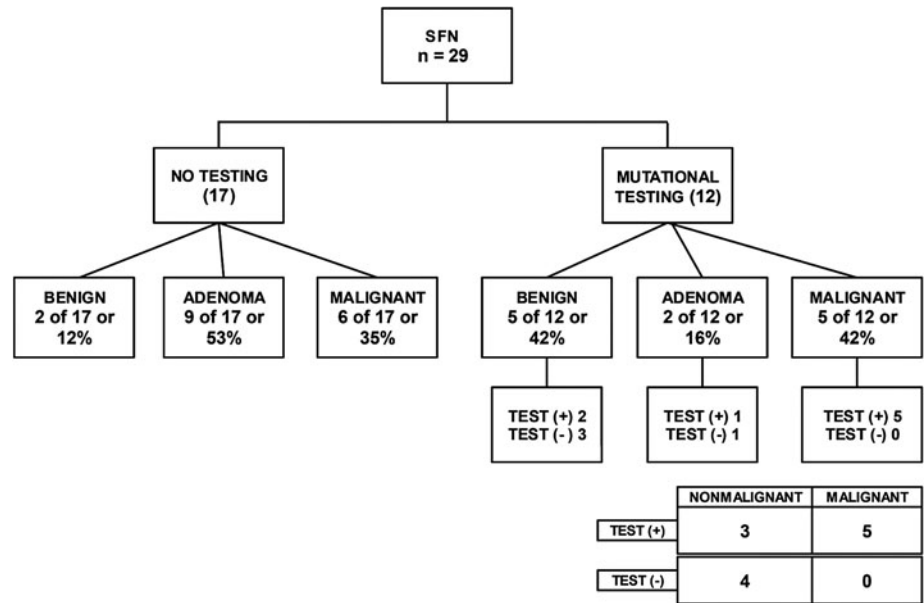


FIG. 1. Surgical histopathology of atypia of undetermined significance (AUS/FLUS) cytology category as a function of whether molecular testing was performed preoperatively. Surgical histopathology diagnosis comparisons between groups with molecular testing versus groups without molecular testing were not statistically significant. Numbers in parenthesis on the 2x2 table show values of seven-gene mutational testing and ThyroSeq<sup>®</sup>, respectively. Sensitivity 85% [confidence interval (CI) 55–98%], specificity 65% [CI 45–81%], PPV 50% [CI 28–72%], NPV 91% [CI 71–99%], positive likelihood ratio (LR) 2.4 [CI 1.4–4.0], and negative LR 0.2 [CI 0.06–0.9].

**FIG. 2.** Surgical histopathology of suspicious for follicular neoplasm (SFN) cytology category as a function of whether molecular testing was performed preoperatively. Surgical histopathology diagnosis comparisons between groups with molecular testing versus groups without molecular testing were not statistically significant. Sensitivity 100% [CI 48–100%], specificity 57% [CI 18–90%], PPV 63% [CI 24–91%], NPV 100% [CI 39–100%], positive LR 2.3 [CI 1.0–5.4], and negative LR 0. All mutational tests were analyzed with Thyro-Seq<sup>®</sup>, and all mutations were RAS mutations.



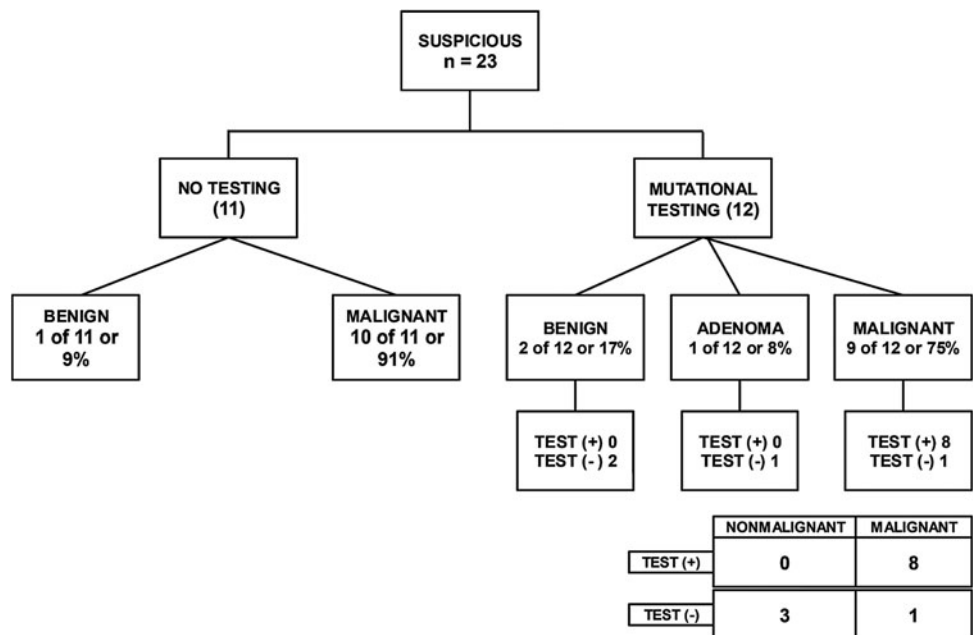
adenoma. Three AUS/FLUS nodules harboring a *PAX8/PPARG* fusion, a *EIF1AX*, or a *TP53* mutation were categorized as follicular adenomas on final histopathology. A *GNAS* mutation was detected in a nodule histologically read as nodular hyperplasia. A total of 14 false-positive results were detected, as defined by the presence of a mutation but a benign histopathology: 11 in the AUS/FLUS and three in the SFN category. Among the 11 AUS/FLUS false positives, seven were read as adenomas, whereas the remaining four were read as nodular hyperplasia. The presence of a RAS mutation accounted for the most false positives (two *HRAS*, six *NRAS*, and one *KRAS*), while a *TP53* mutation and a *PAX8/PPARG* rearrangement were noted in the other two. Similarly, in the FN category, three false positives were noted: one adenoma and two with nodular hyperplasia. Of these, two had a *NRAS* mutation and one had a *HRAS* mutation.

**Discussion**

Initial cytology and mutational analysis of mutational panels were retrospectively correlated with the histopathologic diagnoses of surgeries performed over a two-year period. The data were analyzed relative to whether preoperative FNA molecular testing had been performed. To this end, this study demonstrates surgical histopathologic outcomes reflective of a more generalized medical application with or without molecular testing.

Nearly half of the surgeries were performed after a nodule FNA demonstrated an indeterminate cytology, and more than half of these indeterminate samples were categorized as AUS/FLUS, emphasizing the diagnostic dilemma in this category. The malignancy yield at surgery (i.e., the number with malignancy/number of surgeries) was 48% for the whole cohort.

**FIG. 3.** Surgical histopathology results of suspicious for malignancy (SUSP) category as a function of whether molecular testing was performed preoperatively.



The malignancy yield was 83% for SUSP, 30% and 38% for AUS/FLUS and SFN cytological categories, and 7% for benign cytological diagnosis. Contrary to the idealized malignancy yields as originally described in the BSRTC (3,24), the malignancy detection rate at surgery was higher for AUS/FLUS (30%), as has also been reported in some other studies (25–28). It should be noted that the malignancy rate in this report is based on the selected surgical cohort, and therefore it is higher than in reports based on databases including all biopsied nodules, some of which were not submitted to surgery. As an example, Nikiforov *et al.* reported outcomes of a prospective consecutive sample of 465 AUS/FLUS nodules out of which only 22 were malignant. However, since only 96 of them underwent surgery, the malignancy rate “at surgery” was 22.5% (17). Furthermore, differences in the frequency rates of the different BSRTC categories have been shown to vary across institutions (24,29), emphasizing the need for centers to understand their individual performance rather than to rely solely on idealized BSRTC predictions (22).

Mutational testing was performed in more than half (68/125) of the nodules in the indeterminate cytological category (Table 3). The data clearly show that the mutation panel is more likely to be positive in histologically malignant nodules compared with benign nodules ( $p < 0.0001$ ) or follicular adenomas ( $p = 0.001$ ). The performance of the mutational panel testing in the AUS/FLUS cytology group for sensitivity, specificity, PPV, and NPV was 85%, 65%, 50%, and 91%, respectively (Fig. 1). For the FN group, it was 100%, 57%, 63%, and 100% (Fig. 2). While the data for both the AUS/FLUS and the FN groups had sensitivity and NPV rates on a par with or better than past reports (11,16,17), a significantly lower specificity and PPV were found.

What might account for the lower specificity and PPV found compared with prior reports? False-positive molecular test results (specifically relative to a malignancy diagnosis) could lead to a lower specificity and PPV. Here, the type of mutation affects the estimate of the malignancy rate. The number of mutations detected in the cohort studied here is too low to obtain an accurate estimate of malignancy risks associated with individual mutations. However, *RAS* mutations were most abundant across both the AUS/FLUS and SFN groups, and these mutations were not specific for malignancy (46% malignancy yield rate), and they had a high false-positive rate. A positive *BRAF* mutation correlated with a high risk of cancer on surgical pathology and was mostly detected in the SUSP cytologic category. Also, as has been emphasized in the recent literature, the malignancy prevalence affects the PPV and specificity (22). Had the BSRTC AUS/FLUS and SFN category malignancy prevalence been lower, the specificity and PPV would have been higher.

The expectation for the use of molecular testing preoperatively in the evaluation of thyroid nodules is that it would facilitate management decisions, such as whether to operate or the extent of surgery needed for the indeterminate cytology category of thyroid nodules. Most past studies have compared malignancy rates in mutational panel or gene fusion positive nodules with mutation/fusion-negative nodules but did not compare the malignancy rates in populations who were not tested for mutations (11,16,17). To the authors' knowledge, only one other study has reported data on the surgical histopathologic diagnosis with the use of mutational panel testing in comparison to a group where such molecular

testing was not obtained (21). In that study, the presence of mutations on seven-gene mutational testing led to higher detection of clinically significant thyroid cancer, thereby helping to assess extent of surgery preoperatively. However, the malignancy rate in the mutation-tested group compared to the untested group did not appear to be higher.

In the mutation-positive 22 AUS/FLUS nodules in this study (Fig. 1), 50% were malignant, while untested AUS/FLUS nodules going directly to surgery had a malignancy rate of 31% (9/29). There was no statistical difference of malignancy rates in these two groups ( $p = 0.17$ ). However, for several reasons, this should not be interpreted as a failure of molecular testing to facilitate surgical decision making. Although the risk of malignancy was not different in the tested and untested groups, improvement in malignancy detection with molecular testing is better assessed by the LR than by assessment for statistical significance of these small groups (30,31). Even with a low positive LR of 2.4, the post-test probability increases by nearly 15% (31). Such an increase improves the expected malignancy rate to 20–30% when applied to the 5–15% reported expected AUS category malignancy rates (3). Since past and current recommendations for the SFN cytology group is surgery (3–6), and the BSRTC expected malignancy rate in this group is 15–30%, a positive mutation in an AUS/FLUS nodule is a finding that supports the recommendation for surgery. On the other hand, for mutation-negative nodules, the negative likelihood ratio of 0.24 would reduce the probability of malignancy to a predicted prevalence of <10%, and these nodules could be managed with clinical follow-up. Therefore, collecting an extra sample during initial FNA for a reflex mutational panel testing could avoid repeat FNA and could be both convenient and cost-effective.

Histopathologic assessment is the gold standard for designating a thyroid nodule as hyperplasia, adenoma, or carcinoma. However, follicular lesions of the thyroid are a heterogeneous group, and interpretation is not straightforward. Cases may fall in the morphologic grey zone where a difference of opinion exists even among experts regarding the criteria or threshold of categorizing a particular nodule as non-neoplastic or neoplastic (32–35). It has been proposed that any nodule with a *RAS* mutation should be categorized, at a minimum, as an adenoma rather than a hyperplastic nodule, since the presence of a mutation indicates a proliferating clonal population where the full extent of characteristic morphologic features are yet to be displayed (36). Likewise, it has been proposed that such nodules harboring a *RAS* mutation may be precursors to *RAS*-positive follicular carcinomas or follicular variant papillary carcinomas (36–38), justifying their removal. Perhaps the future of neoplasm or malignancy detection or histopathologic classification will be by mutational status rather than histopathology (8,37). However, of the 18 histologically diagnosed follicular adenomas in this series that were tested with the mutational panel, eight were mutation positive but 10 were not. Of the positive mutations, the majority were *RAS* mutations (6/8 nodules; one nodule had an additional *RET* mutation, and two nodules had *PAX8/PPARG* and *TP53* mutations). Similarly, 6/19 with nodular hyperplasia were positive for a *RAS* mutation. Therefore, to the extent that the surgical histopathologic diagnosis remains the gold standard, and using multiple pathologists is common within a university practice, the mutational panel testing did

TABLE 4. PERFORMANCE COMPARISON OF MUTATIONAL TESTING IN THE PRESENT STUDY TO PUBLISHED LITERATURE

	<i>Present study</i>	<i>Sevem-gene MT</i>	<i>ThyroSeq® v2</i>
<b>Sensitivity</b>			
AUS	85%	63%	91%
FN	100%	57%	90%
<b>Specificity</b>			
AUS	65%	99%	92%
FN	57%	97%	93%
<b>PPV</b>			
AUS	50%	88%	77%
FN	63%	87%	83%
<b>NPV</b>			
AUS	91%	94%	97%
FN	100%	86%	96%
<b>n</b>			
AUS	44	247	98
FN	12	214	143
<b>Malignancy rate</b>			
AUS	30%	14%	23%
FN	38%	27%	27%

Performance comparison of mutational testing in the present study to published literature using the same assays: seven-gene mutational test (MT) (11) and ThyroSeq® v2 (16,17).

not seem to aid in distinguishing a neoplastic (adenoma and malignant) lesion from nodular hyperplasia or other benign diagnoses. However, the presence of a mutation in nodular hyperplasia or follicular adenoma could imply clonal proliferation of cells and possibly a premalignant lesion. Further research on such nodules would provide more insight on this topic.

This study is not without limitations. First, since the number of cases in the subcategories is small (Table 4), it may lack statistical power to demonstrate a difference in the calculated malignancy risks as predicted by the molecular analysis. Second, the malignancy rate is a crude calculation that is heavily dependent on the denominator (total number of surgeries in the AUS/FLUS and SFN categories). The results could be different if everyone who had a FNA that showed AUS/FLUS or SFN underwent surgery. However, it is not possible to achieve this type of ideal scenario in research as well as in a clinical setting. Third, the effect of sample size and the possibility that mutational test-negative nodules were not routinely operated on needs to be considered. In the published literature (11,16,17), the *n* for the true-negative groups was much larger than in the present study. Had a much larger number of mutation-negative nodules been operated on and proven to be histologically benign, thus reducing the pretest malignancy prevalence in the AUS/FLUS group, the positive LR would have risen correspondingly (30). Therefore, the results may underestimate the true positive LR performance of the molecular test. Fourth, second-opinion review of the cytologic and surgical diagnosis by a limited number of expert pathologists has shown reclassification of a significant number of FNA samples (34,39). The data are derived from routine care delivered by eight pathologists; cases were not reanalyzed for histopathologic diagnosis for this study. Finally, the decision for surgery could have been

predetermined by clinical and ultrasound features in both groups, regardless of the results of mutational analysis.

Based on the current literature, it is not clear whether molecular studies need to be performed on initial or repeat FNAs in order to improve the diagnostic yield of malignancy or to reduce the cost of care. A direct comparison with the current standard of care is warranted to demonstrate whether mutational panel testing is superior or non-inferior to repeating FNA. Furthermore, comparison of cost differences between the groups who undergo mutational testing and those who undergo repeat FNA (as measured by type of surgery (21), cost, time, malignancy stage, surgical complications, etc.) may also favor using mutation analysis, and should be investigated in order to establish the utility of specific molecular mutation/gene fusion testing in current practice.

In summary, the present results demonstrate that the molecular panel mutations are more prevalent in the malignant nodules compared with benign or adenomatous nodules. Mutational testing had a high sensitivity and NPV consistent with published data, but a lower than anticipated PPV and specificity. Additional independent studies are needed to evaluate further the utility of molecular testing in the AUS/FLUS and SFN categories.

## Acknowledgments

Research reported in this publication was supported by the National Center for Advancing Translational Sciences of the National Institutes of Health Award Number UL1TR000114. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

## Author Disclosure Statement

No competing financial interests exist.

## References

1. Miller JM, Hamburger JI, Kini S 1979 Diagnosis of thyroid nodules. Use of fine-needle aspiration and needle biopsy. *JAMA* **241**:481–484.
2. Hamberger B, Gharib H, Melton LJ 3rd, Goellner JR, Zinsmeister AR 1982 Fine-needle aspiration biopsy of thyroid nodules. Impact on thyroid practice and cost of care. *Am J Med* **73**:381–384.
3. Cibas ES, Ali SZ, NCI Thyroid FNA State of the Science Conference 2009 The Bethesda System For Reporting Thyroid Cytopathology. *Am J Clin Pathol* **132**:658–665.
4. National Comprehensive Cancer Network 2015. Available at: [www.nccn.org/professionals/physician\\_gls/pdf/thyroid.pdf](http://www.nccn.org/professionals/physician_gls/pdf/thyroid.pdf) (accessed January 8, 2016).
5. Gharib H, Papini E, Paschke R, Duick DS, Valcavi R, Hegedus L, Vitti P, AACE/AME/ETA Task Force on Thyroid Nodules 2010 American Association of Clinical Endocrinologists, Associazione Medici Endocrinologi, and European Thyroid Association Medical guidelines for clinical practice for the diagnosis and management of thyroid nodules: executive summary of recommendations. *Endocr Pract* **16**:468–475.
6. 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 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.
7. Fagin JA, Mitsiades N 2008 Molecular pathology of thyroid cancer: diagnostic and clinical implications. *Best Pract Res Clin Endocrinol Metab* **22**:955–969.
  8. Cancer Genome Atlas Research Network 2014 Integrated genomic characterization of papillary thyroid carcinoma. *Cell* **159**:676–690.
  9. Xing M 2013 Molecular pathogenesis and mechanisms of thyroid cancer. *Nat Rev Cancer* **13**:184–199.
  10. Hodak SP, Rosenthal DS, American Thyroid Association Clinical Affairs Committee 2013 Information for clinicians: commercially available molecular diagnosis testing in the evaluation of thyroid nodule fine-needle aspiration specimens. *Thyroid* **23**:131–134.
  11. Nikiforov YE, Ohori NP, Hodak SP, Carty SE, LeBeau SO, Ferris RL, Yip L, Seethala RR, Tublin ME, Stang MT, Coyne C, Johnson JT, Stewart AF, Nikiforova MN 2011 Impact of mutational testing on the diagnosis and management of patients with cytologically indeterminate thyroid nodules: a prospective analysis of 1056 FNA samples. *J Clin Endocrinol Metab* **96**:3390–3397.
  12. Ohori NP, Nikiforova MN, Schoedel KE, LeBeau SO, Hodak SP, Seethala RR, Carty SE, Ogilvie JB, Yip L, Nikiforov YE 2010 Contribution of molecular testing to thyroid fine-needle aspiration cytology of “follicular lesion of undetermined significance/atypia of undetermined significance.” *Cancer Cytopathol* **118**:17–23.
  13. Cantara S, Capezzone M, Marchisotta S, Capuano S, Busonero G, Toti P, Di Santo A, Caruso G, Carli AF, Brilli L, Montanaro A, Pacini F 2010 Impact of proto-oncogene mutation detection in cytological specimens from thyroid nodules improves the diagnostic accuracy of cytology. *J Clin Endocrinol Metab* **95**:1365–1369.
  14. Nikiforov YE, Steward DL, Robinson-Smith TM, Haugen BR, Klopper JP, Zhu Z, Fagin JA, Falciglia M, Weber K, Nikiforova MN 2009 Molecular testing for mutations in improving the fine-needle aspiration diagnosis of thyroid nodules. *J Clin Endocrinol Metab* **94**:2092–2098.
  15. 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**:E1852–1860.
  16. Nikiforov YE, Carty SE, Chiosea SI, Coyne C, Duvvuri U, Ferris RL, Gooding WE, Hodak SP, LeBeau SO, Ohori NP, Seethala RR, Tublin ME, Yip L, Nikiforova MN 2014 Highly accurate diagnosis of cancer in thyroid nodules with follicular neoplasm/suspicious for a follicular neoplasm cytology by ThyroSeq v2 next-generation sequencing assay. *Cancer* **120**:3627–3634.
  17. 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.
  18. Yip L, Wharry LI, Armstrong MJ, Silbermann A, McCoy KL, Stang MT, Ohori NP, LeBeau SO, Coyne C, Nikiforova MN, Bauman JE, Johnson JT, Tublin ME, Hodak SP, Nikiforov YE, Carty SE 2014 A clinical algorithm for fine-needle aspiration molecular testing effectively guides the appropriate extent of initial thyroidectomy. *Ann Surg* **260**:163–168.
  19. Schuster SC 2008 Next-generation sequencing transforms today’s biology. *Nat Methods* **5**:16–18.
  20. Ferraz C, Eszlinger M, Paschke R 2011 Current state and future perspective of molecular diagnosis of fine-needle aspiration biopsy of thyroid nodules. *J Clin Endocrinol Metab* **96**:2016–2026.
  21. Yip L, Kebebew E, Milas M, Carty SE, Fahey TJ 3rd, Parangi S, Zeiger MA, Nikiforov YE 2010 Summary statement: utility of molecular marker testing in thyroid cancer. *Surgery* **148**:1313–1315.
  22. Ferris RL, Baloch Z, Bernet V, Chen A, Fahey TJ 3rd, Ganly I, Hodak SP, Kebebew E, Patel KN, Shaha A, Steward DL, Tufano RP, Wiseman SM, Carty SE, American Thyroid Association Surgical Affairs Committee 2015 American Thyroid Association statement on surgical application of molecular profiling for thyroid nodules: current impact on perioperative decision making. *Thyroid* **25**: 760–768.
  23. Shrestha RT, Karunamurthy A, Amin K, Nikiforov YE, Caramori ML 2015 Multiple mutations detected preoperatively may predict aggressive behavior of papillary thyroid cancer and guide management—a case report. *Thyroid* **25**:1375–1378.
  24. 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.
  25. Ho AS, Sarti EE, Jain KS, Wang H, Nixon IJ, Shaha AR, Shah JP, Kraus DH, Ghossein R, Fish SA, Wong RJ, Lin O, Morris LG 2014 Malignancy rate in thyroid nodules classified as Bethesda category III (AUS/FLUS). *Thyroid* **24**: 832–839.
  26. Cuhaci N, Arpacı D, Ucler R, Yazgan AK, Kiyak G, Yalcin S, Ersoy PE, Guler G, Ersoy R, Cakir B 2014 Malignancy rate of thyroid nodules defined as follicular lesion of undetermined significance and atypia of undetermined significance in thyroid cytopathology and its relation with ultrasonographic features. *Endocr Pathol* **25**:248–256.
  27. Mathur A, Najafian A, Schneider EB, Zeiger MA, Olson MT 2014 Malignancy risk and reproducibility associated with atypia of undetermined significance on thyroid cytology. *Surgery* **156**:1471–1476; discussion 1476.
  28. Kholova I, Ludvikova M 2014 Thyroid atypia of undetermined significance or follicular lesion of undetermined significance: an indispensable Bethesda 2010 diagnostic category or waste garbage? *Acta Cytol* **58**:319–329.
  29. Fazeli R, Schneider EB, Ali SZ, Zeiger MA, Olson MT 2015 Diagnostic frequency ratios are insufficient to measure laboratory precision with the Bethesda System for Reporting Thyroid Cytopathology. *Acta Cytol* **59**:225–232.
  30. Grimes DA, Schulz KF 2005 Refining clinical diagnosis with likelihood ratios. *Lancet* **365**:1500–1505.
  31. McGee S 2002 Simplifying likelihood ratios. *J Gen Intern Med* **17**:646–649.
  32. Hirokawa M, Carney JA, Goellner JR, DeLellis RA, Hefess CS, Katoh R, Tsujimoto M, Kakudo K 2002 Observer variation of encapsulated follicular lesions of the thyroid gland. *Am J Surg Pathol* **26**:1508–1514.
  33. Baloch ZW, Fleisher S, LiVolsi VA, Gupta PK 2002 Diagnosis of “follicular neoplasm”: a gray zone in thyroid fine-needle aspiration cytology. *Diagn Cytopathol* **26**:41–44.



34. Wang CC, Friedman L, Kennedy GC, Wang H, Kebebew E, Steward DL, Zeiger MA, Westra WH, Wang Y, Khanafshar E, Fellegara G, Rosai J, Livolsi V, Lanman RB 2011 A large multicenter correlation study of thyroid nodule cytopathology and histopathology. *Thyroid* **21**:243–251.
35. Cibas ES, Baloch ZW, Fellegara G, LiVolsi VA, Raab SS, Rosai J, Diggans J, Friedman L, Kennedy GC, Kloos RT, Lanman RB, Mandel SJ, Sindy N, Steward DL, Zeiger MA, Haugen BR, Alexander EK 2013 A prospective assessment defining the limitations of thyroid nodule pathologic evaluation. *Ann Intern Med* **159**:325–332.
36. Bhaijee F, Nikiforov YE 2011 Molecular analysis of thyroid tumors. *Endocr Pathol* **22**:126–133.
37. Gupta N, Dasyam AK, Carty SE, Nikiforova MN, Ohori NP, Armstrong M, Yip L, LeBeau SO, McCoy KL, Coyne C, Stang MT, Johnson J, Ferris RL, Seethala R, Nikiforov YE, Hodak SP 2013 *RAS* mutations in thyroid FNA specimens are highly predictive of predominantly low-risk follicular-pattern cancers. *J Clin Endocrinol Metab* **98**:E914–922.
38. Nikiforova MN, Lynch RA, Biddinger PW, Alexander EK, Dorn GW 2nd, Tallini G, Kroll TG, Nikiforov YE 2003 *RAS* point mutations and *PAX8-PPAR* gamma rearrangement in thyroid tumors: evidence for distinct molecular pathways in thyroid follicular carcinoma. *J Clin Endocrinol Metab* **88**:2318–2326.
39. Gerhard R, Boerner SL 2014 The value of second opinion in thyroid cytology: a review. *Cancer Cytopathol* **122**:611–619.

Address correspondence to:

*Rupendra T. Shrestha, MD*

*Division of Diabetes, Endocrinology and Metabolism*

*University of Minnesota*

*420 Delaware St. SE, MMC 101*

*Minneapolis, MN 55455*

*E-mail: rupen.shrestha@gmail.com*