

Noninvasive Follicular Thyroid Neoplasm with Papillary-Like Nuclear Features: Epidemiology and Long-Term Outcomes in a Strictly Defined Cohort

Rainjade Chung,^{1,2} Haixia Guan,^{3,4} Cecilia Ponchiardi,⁵ Sandra Cerda,⁵ Nitin Marwaha,⁵ Osman H. Yilmaz,⁵ Emma Pinjic,¹ David McAneny,¹ Stephanie L. Lee,⁴ and Frederick Thurston Drake¹

Background: A subset of encapsulated/circumscribed follicular variant of papillary thyroid carcinoma (FVPTC) was reclassified as noninvasive follicular thyroid neoplasm with papillary-like nuclear features (NIFTP) in 2016 to reduce overtreatment of a low-risk tumor. Study objectives were to describe the epidemiology and long-term outcomes of NIFTP in a high-volume, urban, tertiary referral center.

Methods: Among patients enrolled in the Boston Medical Center (BMC) Thyroid Cancer Registry, 110 cases of FVPTC underwent index thyroid surgery at BMC between 2000 and 2016. Historically, BMC pathologists assess all malignant nodules using sections ≤ 0.3 cm with evaluation of the entire nodule and capsule. After review of pathology reports to identify potential NIFTPs, slides were rereviewed using criteria established by the NIFTP Working Group in 2016 and 2018. We evaluated interobserver reliability using Cohen's Kappa coefficient.

Results: Among 110 FVPTCs, 15 (13%) met NIFTP criteria; 11 women and 4 men, age range 31–64 (mean 47.5) years. Mean tumor diameter was 1.7 cm (compared with 2.2 cm for FVPTC). Among NIFTP cases, there were no lymph node metastases, distant metastases, or tumor recurrences. All NIFTP cases were American Thyroid Association (ATA) low risk compared with only 68% of FVPTC ($p=0.011$). Among FVPTCs, 14% had positive lymph nodes at index operation. Four patients (4%) had distant metastases. Mean follow-up time was 46 and 69 months for FVPTC and NIFTP, respectively. Among FVPTCs with an excellent response to therapy (2015 ATA guidelines), there were no recurrences. Just over half ($n=8$) of patients with NIFTP received postoperative radioactive iodine (RAI) therapy. Concordance between pathologists was high for ruling out NIFTP (75%), but only 36% for ruling in NIFTP. Overall, for NIFTP designation, Cohen's Kappa was 0.39, which is considered fair.

Conclusions: Although this is a relatively small cohort, all NIFTP specimens underwent updated pathology review consistent with current guidelines; mean follow-up was nearly 6 years. NIFTP represents a small fraction of the total papillary neoplasia diagnosed at this tertiary referral center (2.3%). None of the NIFTP cohort experienced an adverse oncologic event, and there were no regional or distant metastases. Over 50% of patients with NIFTP received RAI. Thus, the NIFTP reclassification may substantially reduce the number of patients who require adjuvant therapies, such as completion surgery or RAI.

Keywords: NIFTP, follicular variant of papillary thyroid cancer, incidence, recurrence, *BRAF*^{V600E} mutation, radioactive iodine therapy

¹Department of Surgery, Section of Endocrine Surgery, Boston University, Boston, Massachusetts, USA.

²Department of Surgery, Columbia University Medical Center, New York, New York, USA.

³Department of Endocrinology and Metabolism, The First Hospital of China Medical University, Shenyang, China.

⁴Department of Medicine, Section of Endocrinology, Nutrition, Diabetes, and Weight Management, Boston University, Boston, Massachusetts, USA.

⁵Department of Pathology, Boston University, Boston, Massachusetts, USA.

Introduction

NONINVASIVE ENCAPSULATED FOLLICULAR variant of papillary thyroid cancer (FVPTC) has long been considered clinically distinct from the infiltrative or invasive follicular variants of PTC, with a lower risk profile for recurrence and metastatic disease (1). This distinction was codified in a 2016 article by Nikiforov *et al.* that put forth classification criteria and proposed a new nomenclature for this entity: noninvasive follicular thyroid neoplasm with papillary-like nuclear features (NIFTP) (2). The work of this multinational, multidisciplinary group showed that, in 109 patients diagnosed with NIFTP and a median follow-up of 13 years, there was no recurrence or evidence of distant disease. Ultimately, the authors proposed reclassifying NIFTP from a malignant lesion to a neoplasm with low (or even no) malignant potential to reduce the stigma and psychological toll of a “cancer” diagnosis as well as resulting overtreatment.

Several other groups have demonstrated similar results supporting this reclassification (1,3,4), but some studies have suggested that a more cautious approach be taken regarding NIFTP (5–7) because positive regional lymph nodes and even distant metastases were found in their NIFTP cohorts. Moreover, interobserver reliability is a challenge to consistent and accurate diagnosis, particularly as some of the NIFTP criteria may be subjective, such as assessing percent solid growth and the meaning of “well-formed” papillae. Most published series have been small, between 50 and 100 cases; to our knowledge, the largest series of NIFTP cases with complete pathologic review and long-term follow-up comprised 129 cases (3). Interestingly, a recently published editorial by the “NIFTP Working Group” found that many—although not all—of the cases published in the literature demonstrating regional lymph node metastases likely did not meet “strict” criteria for NIFTP when the individual case was retrospectively rereviewed (8).

Papillary thyroid cancer (PTC) is by far the most common thyroid cancer and its incidence is increasing (9). FVPTC is responsible for 10–25% of all PTC and has been increasing, both in incidence and as a proportion of overall PTC (2,10). Thus, outcomes and treatment recommendations for this particular segment of thyroid cancer patients is of interest not just to patients and clinicians, but also to the entire health systems, private payers such as insurance companies, and social welfare entities such as Medicare and Medicaid. Given its high incidence, reducing overtreatment of even a small segment of FVPTC patients could be of significant benefit to individuals, health systems, and payers alike.

It is well documented that there is high interobserver variability in diagnosing FVPTCs (11–13). Distinguishing NIFTP from FVPTC may also be challenging, an observation underscored by the fact that universal agreement was not achieved even by the expert thyroid pathologists who collaborated in the 2016 project establishing the guidelines for NIFTP. In some cases, their assessments differed as to whether a lesion was a classic PTC, an infiltrative FVPTC, an invasive encapsulated FVPTC, an NIFTP, or a follicular adenoma. Interestingly, the authors noted that it was the degree of nuclear features of PTC that correlated most closely with agreement between pathologists regarding how to classify a tumor (2).

Based on these recent classification changes in thyroid pathology and because of the conflicting data among patient series published to date, this research study was designed with five goals: (i) to establish the epidemiology of NIFTP among PTC patients at an urban, socioeconomically diverse, tertiary referral center, (ii) to compare the frequency of metastases and response to therapy for patients with NIFTP compared with those with infiltrative or invasive FVPTC, (iii) to compare the frequency of the *BRAF*^{V600E} mutation in our cohort of NIFTP and FVPTC patients, (iv) to study the proportions of pathologic findings that excluded FVPTC specimens from meeting NIFTP criteria, and (v) to assess interobserver variability in NIFTP classification, with a particular emphasis on the presence and proportion of papillae and percent solid component.

We had a special opportunity to conduct this study because the Department of Pathology at the Boston University School of Medicine has a long-standing practice of submitting all tumors entirely for histologic evaluation in sections that are 0.3 cm or less in thickness with evaluation of the entire capsule and the entire nodule. This is notable, because some previous studies may have relied on review of specimens that were not prepared appropriately to rule in or rule out NIFTP (i.e., they did not have three sections per cm, with evaluation of the entire capsule and entire tumor) (8). Although this series is not large, we have confidence that the NIFTP designation is in full compliance with both the 2016 criteria and 2018 update. Because the Boston Medical Center (BMC) health system is both the primary care home and the tertiary referral center for a large and long-term population, this study is strengthened by robust longitudinal follow-up. BMC cares for a diverse patient population and this cohort of patients may reflect different epidemiologic patterns than other studies of NIFTP. This is another important aspect of the current project since widely varying rates of NIFTP have been reported. This observational study was designed and carried out in accordance with the principles articulated in the STROBE statement (14).

Materials and Methods

Study type

This is a retrospective cohort from a high-volume, urban, tertiary referral center in North America where a diverse patient population receives both primary and subspecialty care.

Data source

The Boston University/BMC Thyroid Cancer Registry (ThyroCARE) is an IRB-approved cancer registry that retrospectively enrolled former and existing thyroid cancer patients and now prospectively enrolls new thyroid cancer patients (including all those with NIFTPs). Patients are included in the registry if they had any aspect of their thyroid cancer care at BMC. In addition to data currently contained in ThyroCARE, we utilized the BMC electronic medical record for additional data as necessary. We had access to the original pathology specimens (which were rereviewed) for all patients included in the study.

Subject inclusion criteria

To be included in this study cohort, the patients had to have undergone their index thyroid cancer operation at BMC, and

the specimen had to be prepared according to our institutional specifications, three sections/cm and evaluation of the entire capsule and nodule for all malignant lesions. All surgical pathology reports included information about tumor architecture and cytology, tumor classification, mutational analysis for *BRAF*^{V600E}, capsule invasion, local invasion, lymphatic spread, and angioinvasion. Many patients in the ThyroCARE database had their index operations at other institutions, and those patients were not eligible for inclusion in this particular study because their pathology specimens were not prepared according to our institutional protocols and specifications.

The ThyroCARE database was queried for all patients with an FVPTC or NIFTP and whose operation was performed at BMC. This was done via a keyword search for “follicular.” Follicular thyroid cancers were excluded, and FVPTC or NIFTP pathology reports were then screened by the senior authors (S.L.L., F.T.D.) to identify NIFTP candidates. Cases were only included if researchers had access to the original pathology slides.

Pathology rereview

After review of pathology reports to identify NIFTP candidates, all specimens were rereviewed by the BMC pathology team (S.C., C.P.) using the inclusion criteria established by Nikiforov and others and revised in the 2018 publication: (7,8) encapsulation or clear demarcation, follicular growth pattern, no papillae, no psammoma bodies, <30% solid/trabecular/insular growth pattern, nuclear score 2–3, no vascular or capsular invasion, no tumor necrosis, no high mitotic activity, and no *BRAF*^{V600E} mutation (if testing was done). In this study, if both FVPTC and NIFTP were present simultaneously, the patient was considered part of the FVPTC group. In addition, no patients were categorized into the NIFTP group if they had concomitant classic PTC or any other thyroid malignancy. For any patients with multifocal disease in the NIFTP cohort, all lesions had to meet the NIFTP criteria. Both infiltrative FVPTC and invasive encapsulated FVPTC were included in the FVPTC cohort in this study, without stratification.

Variables included

The variables included were sex, age, race/ethnicity, tumor characteristics on final pathology (size, encapsulation, lymphovascular invasion, etc.), American Thyroid Association (ATA) initial risk classification, and ATA response to therapy. Chart review was conducted by trained abstractors to input retrospective data into ThyroCARE. Information on initial ATA risk stratification and follow-up information regarding biochemical or structural recurrence are included in the registry. Mean follow-up time, which is a strength of this study because it is precisely defined, was based on the date of the patients’ most recent encounter in the BMC medical record.

Outcomes assessed

The outcomes assessed were presence of metastases at index operation, initial ATA risk stratification, response to therapy (acceptable response, indeterminate, biochemically incomplete, structurally incomplete), and cancer-specific mortality.

Interobserver variability and reproducibility of results

After the original pathology review was performed by the department’s most senior thyroid pathologist, a subset of 32 patients was selected for rereview by two pathologists blinded to the original designation. The same tumor blocks and slides were utilized as in the initial evaluation. Eight NIFTPs were randomly selected from the NIFTP cohort and 24 FVPTCs were randomly selected, although we purposefully selected from a subset of cases with >1% papillae ($n=8$), cases with 1 or more papillae but <1% overall ($n=16$), or >30% solid ($n=12$) to create the non-NIFTP cohort for secondary analysis. Other than the presence of papillae or proportion of solid growth, all 24 non-NIFTP lesions met the remainder of the histologic criteria for NIFTP. Our objective for this subanalysis was to evaluate the reproducibility of an NIFTP diagnosis and, in particular, the reproducibility of assessments for papillary structures and degree of solid components.

Statistical analysis

Categorical variables were compared via the Fisher exact test, and continuous variables by Student’s *t*-test. A result was considered statistically significant if $p < 0.05$. Because the cohort of patients that met all the NIFTP criteria was only 15 patients, it was not appropriate to utilize multivariable methodologies to compare long-term outcomes with FVPTCs. For interobserver reproducibility, we utilized simple percentages of concordance between the first and second pathologic assessment. To address concordance that occurs due to chance, we further deployed Cohen’s Kappa coefficient to quantify interobserver reliability (15).

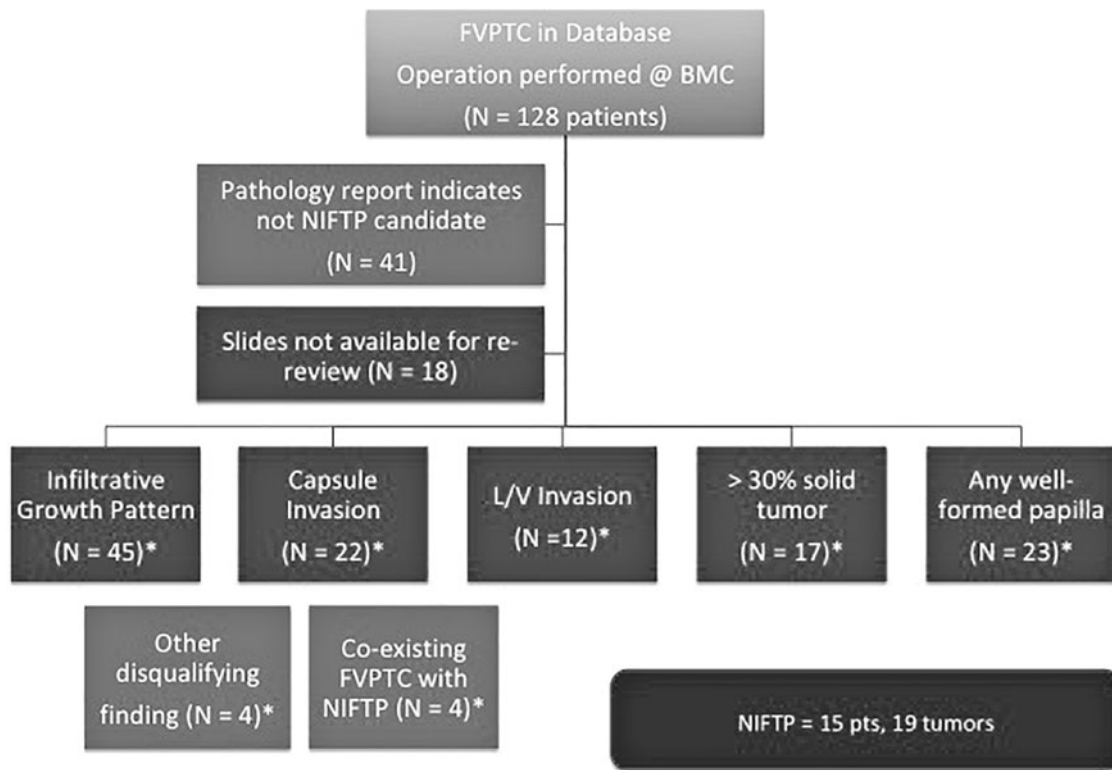
Human subjects’ approval

The Boston University School of Medicine Institutional Review Board approved this study.

Results

Of 1212 patients included in the ThyroCARE database from 2000 to 2016, a complete pathology report was available for 765 patients at the time of this analysis (Summer 2018). Initial screening of pathology reports identified 128 FVPTCs and/or NIFTP candidates from 663 total PTC cases between 2000 and 2016. Of these 128 cases, 18 did not have slides available for rereview and were thus excluded from the study (Fig. 1). As is well described, patients frequently have multiple foci of cancer within a single specimen. Within this group of 765 patients with full pathology reports, we observed the following breakdown of neoplasms and cancers (each patient can have up to five cancer foci included in the ThyroCARE cancer registry): there were 2229 total neoplastic foci, including 1629 PTCs (73.1%), 464 (20.8%) FVPTCs (including NIFTPs), 45 follicular carcinomas (2%), 36 (1.6%) Hurthle cell carcinomas, 22 (1%) medullary thyroid carcinomas, 24 (1.1%) poorly differentiated cancers, and 9 (0.4%) anaplastic thyroid cancers.

Of the 110 patients with FVPTC or NIFTP conclusively identified from the database, 15 (13%) patients had tumors that met criteria for NIFTP (2.3% of all PTCs). Overall, the distribution of age, sex, and race/ethnicity was similar for both patients with NIFTP and patients with FVPTC. Both



* Pathology specimens could have more than one finding

FIG. 1. Allocation of patients to the FVPTC or NIFTP study arms. This figure illustrates the flow of patients from initial identification in the ThyroCARE database to final allocation into one of the two study designations, and it also summarizes the frequency with which various NIFTP criteria were not met among the FVPTC cohort. FVPTC, follicular variant of papillary thyroid carcinoma; NIFTP, noninvasive follicular thyroid neoplasm with papillary-like nuclear features.

cohorts were predominantly female (FVPTC 78% vs. NIFTP 73%) with a mean age 48 and 49 years, respectively. The proportion of Caucasian patients in the FVPTC group was 49% versus 66% in the NIFTP group ($p=0.25$) (Table 1). Four patients had coexisting FVPTC and NIFTP and these were considered part of the FVPTC cohort per original study design (Fig. 1).

Disease and treatment characteristics

Unifocal lesions were most common (79% total). Among FVPTCs, 16% (15 of 95) had multiple tumors, and 8% had bilateral tumors (8 of 95, including one in a pyramidal lobe). For NIFTPs, those percentages were 27% (4 of 15) and 20% (3 of 15), respectively. The mean diameter of FVPTCs was slightly larger than NIFTPs (2.2 cm vs. 1.7 cm) although this was not statistically significant (Table 1).

Among the FVPTC patients, a majority (96%) of patients had total/completion thyroidectomy ($n=92$) with three patients receiving lobectomy alone. At diagnosis, 14% of FVPTC patients had positive lymph nodes at their initial operation and 4% had distant metastases. Of the 43 FVPTC cases tested for *BRAF*^{V600E} mutations, 8 were positive (18.6%). Two-thirds of FVPTCs (68%) were classified as ATA initial low risk, with 5% classified as high risk. Nearly two-thirds of the FVPTC patients were treated with radioactive iodine (RAI) therapy. A majority of FVPTC patients

(92%) had an acceptable response to therapy as per the ATA 2015 guidelines. However, 6% had a biochemically incomplete response and 6% had a structurally incomplete response.

The NIFTP population stood in contrast to the FVPTC population in that there were no regional lymph node metastases nor distant metastases at initial diagnosis. All of the NIFTP patients were ATA low risk (compared with only 68% of FVPTC patients, $p=0.011$), and 100% had an excellent response to therapy (compared with 81% of FVPTC patients, $p=NS$). Of the six NIFTP cases tested for *BRAF*^{V600E} mutations, all were negative. Among NIFTP patients, lobectomy alone ($n=1$) or total thyroidectomy ($n=12$) was the initial treatment, with three patients undergoing central node dissection. Two patients underwent completion thyroidectomy. Approximately half of the NIFTP patients ($n=8$) received postoperative RAI therapy. In long-term follow-up (mean follow-up time was nearly 6 years [70 months]), NIFTP patients had no lymph node metastases, distant metastases, or tumor recurrences.

Cancer-specific outcomes

All patients with NIFTP were classified as ATA low risk after surgery, and there have been no recurrences (biochemical or structural). Among the FVPTCs with an excellent response to therapy, there have been no recurrences; however, among

TABLE 1. DEMOGRAPHICS AND TUMOR CHARACTERISTICS

<i>n</i>	<i>FVPTC</i>	<i>NIFTP</i>	<i>p</i>
	<i>n</i> = 95, foci = 113	<i>n</i> = 15, foci = 19	
Male, <i>n</i> (%)	21 (22)	4 (27)	NS
Mean age (years)	48 (SD 16)	49 (SD 11)	NS
Race/ethnicity, <i>n</i> (%)			NS
White	48 (49)	10 (66)	
Black	23 (24)	3 (20)	
Hispanic	13 (14)	1 (7)	
Asian	9 (10)	1 (7)	
Other/unknown	2 (2)	0	
Mean diameter (cm)	2.2 (SD 1.7)	1.7 (SD 0.85)	NS
Total thyroidectomy	96% (88 TT, 4 CT)	93% (12 TT, 2 CT)	NS
Lobectomy, <i>n</i> (%)	3 (4)	1 (7)	NS
Node dissection, <i>n</i> (%)			NS
Central node	39 (42)	3 (20)	NS*
Lateral node	8 (8)	0	
No dissection	48 (50)	12 (80)	
Metastases at initial diagnosis (total)	19	0	NS*
Central nodes, <i>n</i> (%)	6 (6)	0	
Lateral nodes, <i>n</i> (%)	8 (8)	0	
Distant, <i>n</i> (%)	4 (4)	0	
	1 bone/3 lung		
Multifocal, <i>n</i> (%)	15 (16)	4 (27)	NS
<i>BRAF</i> ^{V600E} positive	19% (43 tested, 8 positive)	0% (6 tested, 0 positive)	NS

The presence of a microcarcinoma did not qualify a patient as having a multifocal process, which was defined in this analysis as the presence of an additional cancer (or NIFTP) ≥ 1.0 cm.

*For nodal dissections and presence of metastases, the Fisher exact tests were calculated for any node dissection vs. no node dissection and for any metastases vs. no metastases.

CT, completion thyroidectomy; FVPTC, follicular variant of papillary thyroid carcinoma; NIFTP, noninvasive follicular thyroid neoplasm with papillary-like nuclear features; NS, not significant; SD, standard deviation; TT, total thyroidectomy.

those with an indeterminate or incomplete response, there were 4 distant metastases and 14 patients had central and/or lateral neck involvement. One patient with FVPTC has required reoperation for a recurrent lymph node (level 4) in the setting of a prior modified radical neck dissection. No patients with NIFTP have required reoperation for recurrence. Mean follow-up time was 46 and 70 months for FVPTCs and NIFTPs, respectively. There were no cancer-specific deaths in either group (Table 2).

NIFTP criteria

Eighty-seven tumors were categorized as NIFTP candidates upon review of the initial pathology reports. All of these lesions underwent repeat pathology review by two coauthors (S.C., C.P.). Based on this updated pathology review, we tabulated the criteria that the tumors violated that ruled them out as NIFTPs (Fig. 1). The most common exclusion criterion was an infiltrative growth pattern (52%). Capsular invasion (25%) and ≥ 1 papilla (26%) were the next most frequent criteria; the latter finding is particularly notable since this is one of the criteria that were modified in the 2018 criteria revision (7). Extrathyroidal extension (6%) was the least common feature of these tumors that ruled out NIFTP. Fourteen percent of these tumors demonstrated lymphovascular invasion. Lesions could exhibit more than one of the exclusion criteria, which is why the percentages add up to $>100\%$.

Of the 23 specimens with any papillae, there were 19 specimens with $<1\%$ overall, but 1 or more well-formed

papillae. There were 11 patients for whom this was the only exclusion criteria using the updated 2018 criteria. This has been an area of change in terms of defining NIFTP, and we looked at these 11 cases separately (note that our NIFTP cohort does not include these 11 patients). Six of these 11 patients had RAI treatment. All 11 patients were ATA initial stage low risk. Eight were judged to have excellent response to therapy and there is insufficient follow-up for three patients to make that designation. Among five of these cancers that underwent *BRAF*^{V600E} testing, two were positive (40%), compared with zero among the six NIFTP cases that met the strict inclusion criteria and underwent *BRAF*^{V600E} testing.

Interobserver reproducibility

Concordance between pathologists was high for ruling out NIFTP (75%), but only 36% for ruling in NIFTP. Overall, for NIFTP designation, Cohen's Kappa (κ) coefficient for interobserver reliability was 0.39, which is considered fair. We were specifically interested in two of the more subjective (16) components of the NIFTP criteria: presence and proportion of papillary structures and whether lesions had less than or greater than 30% solid growth. The two pathology assessments were concordant 14/32 times (44%) in terms of $<1\%$ papillae or presence/absence of a single papilla (for $<1\%$ papilla, $\kappa = -0.06$, and for no papilla, $\kappa = -0.13$). Cohen's Kappa is interpreted on a -1.0 to 1.0 scale; thus, our results for papilla suggest disagreement that is slightly beyond simple chance

TABLE 2. ADJUVANT THERAPIES, INITIAL RISK STRATIFICATION, AND LONG-TERM OUTCOMES

	FVPTC	NIFTP	P
Treated with RAI, <i>n</i> (%)	62 (65)	8 (53)	NS
Mean follow-up time (months)	46.2	69.8	
ATA initial risk, <i>n</i> (%)			0.011
Low risk	67 (68)	15 (100)	
Intermediate risk	23 (24)	0	
High risk	5 (5)	0	
Response to therapy (based on most recent evaluation), <i>n</i> (%)			
Excellent response	77 (81)	15 (100)	NS (0.13)
Indeterminate response	6 (6)	0	
Biochemical incomplete	6 (6)	0	
Structural incomplete	6 (6)	0	
Local lymph nodes	2 ^a (2)	0	
Distant metastases	5 (5)	0	
Recurrence (among those with acceptable response to therapy)	0	0	NS
Cancer-specific deaths	0	0	NS

Bold indicates $p < 0.05$.

^aOne patient with a positive lymph node on whole-body iodine scan and an incomplete response to treatment also had distant metastases; thus, this patient is counted twice. The other five patients with structural incomplete responses have distant metastases only (one bone, four lungs). Of note, three patients had positive LNs on initial post-treatment scan, but over the follow-up period, developed thyroglobulin levels that meet criteria for “excellent response” and thus are counted in that category. The significant p -value for the comparison of ATA initial risk is for the groups: low risk versus intermediate+high risk combined.

ATA, American Thyroid Association; LN, lymph nodes; RAI, radioactive iodine.

($\kappa = 0.0$ indicates agreement that is no better than chance). Because of purposeful sampling, this cohort of patients had a high proportion of lesions identified by the senior thyroid pathologist as having few (i.e., <1%) or no papillary structures: 24/32 (75%) and 16/32 (50%) of the time, respectively. By contrast, this assessment was made in only 14 of 32 cases (44%) by the junior pathologist (both for <1 and <1%). Concordance was significantly higher for <30% solid at 22/32 (69%), and this finding had the highest interobserver reliability ($\kappa = 0.41$), which is considered moderate.

Of the four original NIFTP cases that were judged to be FVPTCs on a blinded second pathology review, three of four had a disagreement related to the presence of papilla and three of four had a disagreement related to percent solid growth; these features were the only disagreement out of all NIFTP criteria. On the secondary pathology review, there were three cases judged to be NIFTPs that had been called FVPTCs on the original review. Thus, there were seven total NIFTP disagreements. No disagreements resulted from an assessment that there were <1% papillae but at least one well-formed papilla. In other words, the 2018 change in NIFTP criteria would not have impacted the seven disagreements identified between the two pathology reviews.

Finally, related to the three cases identified as NIFTPs on secondary review that were not originally considered NIFTPs: although these cases did not meet criteria as judged by the senior thyroid pathologist, they had the same oncologic char-

acteristics and outcomes as the 15 patients in the original NIFTP cohort: no positive lymph nodes, *BRAF*^{V600E} negative, no metastases, low or undetectable thyroglobulin stable ultrasound, and no recurrences at a median follow-up of 6 years.

Discussion

The BMC Pathology Department’s long-standing practice of consistently reviewing the entirety of the capsule and entire nodule in sections that are ≤ 0.3 cm in thickness established our ability to conduct a long-term retrospective review of tumors that can retroactively and reliably be identified as NIFTPs. When careful pathologic evaluation is performed, NIFTP represents a small fraction of the total PTCs diagnosed at this institution (2.3%). The actual number of NIFTPs is likely a conservative one given that we did not have slides available for rereview in 18 patients who were potential NIFTP candidates, and we excluded patients who did not have their original operation at BMC, which may lead to some referral bias. However, several studies have reported NIFTP incidences of around 2% (5,6,17), consistent with our series. This is important because earlier studies suggested that it might be as high as 15–25% (1–3,8,10).

Some of these differences are likely related to how carefully the entire tumor is routinely examined, how strictly the most recent exclusion criteria have been applied in other series, and, as per Kakudo *et al.*, differences may also be related to how likely pathologists are to categorize lesions with mild nuclear changes as either follicular adenoma or FVPTC or NIFTP (8). As illustrated by our small, subanalysis of interobserver reliability, subjectivity of certain histologic criteria likely also influences incidence. Our study is limited by small numbers, but the data presented here suggest that minor differences in the assessment of a “well-formed” papilla or the percent solid growth may have little influence on oncologic outcomes, particularly when all other criteria are met.

Although the incidence of NIFTP may be low when strict criteria are applied, this diagnosis can decidedly impact patient care and clinical decision-making for specific patients. In this cohort of 15 patients, with respect to adjuvant therapies provided after the index operation, 2 patients would have avoided completion surgery, 8 (50%) would have avoided RAI therapy, and all would have avoided a cancer diagnosis. With wider adoption of the 2015 ATA guidelines, lobectomy is becoming more prevalent as initial therapy for solitary nodules with indeterminate cytology (18). A diagnosis of NIFTP on final surgical pathology should further increase the confidence of the surgeon, endocrinologist, and patient that there is no need for completion thyroidectomy or RAI treatment.

Moreover, using careful pathologic evaluation and strict application of the NIFTP criteria, the patients in our cohort had excellent outcomes. None of the patients with NIFTP in this cohort had local or distant metastatic disease, and all of them had an excellent response to therapy. This stands in contrast to the local and metastatic disease in patients with FVPTC in our study and to other studies in which the authors found metastatic disease in their NIFTP cohorts. For example, Parente *et al.*, who reported the use of rigorous morphologic criteria to identify NIFTPs, found five patients with nodal metastases and one distant metastasis (lung) over a mean follow-up of 5.7 years (5). Kim *et al.* had 9 patients among 74 NIFTPs with positive central neck lymph nodes; however, over half of these patients had concomitant classic PTCs. Cho *et al.* also

identified two patients with central lymph node metastases but no extraregional metastases over a median follow-up of 37 months (6). Of important note, both of these cases were re-reviewed and demonstrated “florid” nuclear features of PTC, prompting the recommendation that potential NIFTPs with nuclear scores of three should undergo a detailed search of the entire tumor for papillae and consideration of $BRAF^{V600E}$ testing. Finally, NIFTP specimens in this cohort meeting the strict NIFTP criteria had no $BRAF^{V600E}$ mutations, but among those FVPTC specimens tested, nearly 20% were $BRAF^{V600E}$ mutation positive. Perhaps most notably, two of five (40%) cancers that met all NIFTP criteria except that they had at least one well-formed papilla (but still <1% papillae) were positive for a $BRAF^{V600E}$ mutation.

These findings differ from the study by Xu *et al.* published in 2019 (16), in which lesions with <1% papillae had no nodal metastases and, among the 24 who were tested, only one case was $BRAF^{V600E}$ positive (this case also had 10% tall cell component, so would not have met NIFTP criteria). However, our results are consistent with a recent editorial from the NIFTP Working Group and data published by Cho *et al.* in which two NIFTP cohorts were compared, one with no papillae and one with <1% papillae (6,8). In the latter study, there were no $BRAF^{V600E}$ mutations in the “no papillae” group, but 10% of tumors in the <1% papillae group carried a $BRAF^{V600E}$ mutation. Moreover, in the study by Kim *et al.*, when they restricted the definition of NIFTP to no papillae, their number of patients with $BRAF^{V600E}$ dropped from nine to one (and possibly zero), as the authors suspected that the $BRAF^{V600E}$ mutation in the remaining patient was from a concomitant classic PTC rather than the NIFTP.

Thus, based on our data and other published series, we agree that the presence of a $BRAF^{V600E}$ mutation should be an exclusion criterion for NIFTP, as should the presence of a single well-formed papilla. Surgeons and endocrinologists must continue to follow the outcomes of these tumors over the long term, as we are doing with ThyroCARE, and this must include data on specific NIFTP histologic criteria as well as data from mutational analyses; these data may clarify whether NIFTP criteria need to evolve further. As proposed by the Memorial Sloan Kettering group, this could even lead back to a *less* restrictive set of criteria, where, for example, <1% papillae is again acceptable for NIFTP instead of the current “no papilla” criterion (16). Moreover, such prospectively accumulated data sets should include preoperative information such as ultrasound characteristics, cytology features, and mutational analyses so that we can develop the capacity to better identify these lesions before surgery and guide patients toward less extensive surgery at their initial operation.

This study has certain limitations, particularly the small number of NIFTPs. This small cohort may not be representative of the NIFTP population, and we may be underestimating the rate of adverse oncologic events. However, this study also has several significant strengths: first, NIFTP was very strictly defined according to the 2018 criteria and all potential NIFTPs were hand-reviewed by a senior, experienced thyroid pathologist to ensure that they met all criteria; second, none of the included NIFTP cases had any concomitant thyroid malignancies, including micro-PTCs; finally, much of the care provided at BMC is in long-standing clinical care relationships and our follow-up is both comprehensive and long term (and prospectively documented in an IRB-approved Redcap© database).

Observational cohort studies of this new entity—NIFTP—are critical so that our understanding of their epidemiology and behavior can be expanded even as this characterization is already starting to influence clinical decision-making. This shift in decision-making is critical as it would essentially change a cancer diagnosis to a premalignant neoplasm, or even a neoplasm with no malignant potential. However, there is an important distinction between preventing overtreatment and discouraging follow-up and monitoring. Until more data and guidelines are published regarding the long-term outcomes of NIFTP, clinicians should continue to encourage follow-up and surveillance of these patients for recurrence or metastatic disease, particularly given the variability in consistently identifying NIFTPs versus FVPTCs, which has been shown both in other studies and by the fair interobserver reliability in the current study.

Indeed, we recommend that institutions develop their own strict protocols for identifying and confirming the diagnosis of NIFTP, that the 2018 criteria be strictly applied, and that a standardized reporting template be implemented so that all criteria are consistently assessed; we recommend that all NIFTPs be reviewed in an expert multidisciplinary conference before any decisions about adjuvant therapy; we recommend that outcomes be tracked prospectively; and finally, we recommend that testing for $BRAF^{V600E}$ be considered for any potential NIFTP if not done preoperatively on a fine needle aspiration specimen, as this is also an objective assessment that may be less susceptible to interobserver variation.

In conclusion, we present a cohort of patients treated at an urban, academic, tertiary care center with a diagnosis of NIFTP based on the 2018 criteria. The incidence of NIFTP among PTC patients at this center was 2.3%. In this cohort of patients, strictly defined NIFTPs represent indolent lesions that have demonstrated no malignant potential over a mean follow-up time of nearly 6 years. Similar to other recent publications, some tumors that meet all the NIFTP criteria except that they contained a single well-formed papilla were associated with the presence of a $BRAF^{V600E}$ mutation but, among those NIFTPs that had absolutely no papillae, the prevalence of $BRAF^{V600E}$ was zero. These data support the hypothesis that NIFTP patients have an excellent prognosis. The data further support the concept that completion thyroidectomy and adjunctive RAI treatment are unnecessary, and our current clinical practice is in line with this. Through prospective enrollment in ThyroCARE, we will be able to study longer term outcomes, and such prospective methodologies will be necessary to study NIFTP in greater detail with more power to ensure that this treatment paradigm maintains good results. High-volume institutions should consider collaborating on a prospective registry of neoplasms that meet strict NIFTP criteria to allow long-term follow-up.

Author Disclosure Statement

No competing financial interests exist.

Funding Information

Development of the project’s REDCAP database was supported by the Boston University Clinical & Translational Science Institute (CTSI), grant number: 1UL1TR001430.

References

1. Thompson LD 2016 Ninety-four cases of encapsulated follicular variant of papillary thyroid carcinoma: a name change to Noninvasive Follicular Thyroid Neoplasm with Papillary-like Nuclear Features would help prevent overtreatment. *Mod Pathol* **29**:698–707.
2. Nikiforov YE, Seethala RR, Tallini G, Baloch ZW, Basolo F, Thompson LDR, Barletta JA, Wenig BM, Ghuzlan A Al, Kakudo K, Giordano TJ, Alves VA, Khanafshar E, Asa SL, El-Naggar AK, Gooding WE, Hodak SP, Lloyd RV, Maytal G, Mete O, Nikiforova MN, Nosé V, Papotti M, Poller DN, Sadow PM, Tischler AS, Michael TR, Wall KB, LiVolsi VA, Randolph GW, Ghossein RA 2016 Nomenclature revision for encapsulated follicular variant of papillary thyroid carcinoma a paradigm shift to reduce overtreatment of indolent tumors. *JAMA Oncol* **2**:1023–1029.
3. Rosario PW, Mourão GF, Nunes MB, Nunes MS, Calsolari MR 2016 Noninvasive follicular thyroid neoplasm with papillary-like nuclear features. *Endocr Relat Cancer* **23**: 893–897.
4. Xu B, Tallini G, Scognamiglio T, Roman BR, Tuttle RM, Ghossein RA 2017 Outcome of large noninvasive follicular thyroid neoplasm with papillary-like nuclear features. *Thyroid* **27**:512–517.
5. Parente DN, Kluijfhout WP, Bongers PJ, Verzijl R, Devon KM, Rotstein LE, Goldstein DP, Asa SL, Mete O, Pasternak JD 2018 Clinical safety of renaming encapsulated follicular variant of papillary thyroid carcinoma: is NIFTP truly benign? *World J Surg* **42**:321–326.
6. Cho U, Mete O, Kim MH, Bae JS, Jung CK 2017 Molecular correlates and rate of lymph node metastasis of non-invasive follicular thyroid neoplasm with papillary-like nuclear features and invasive follicular variant papillary thyroid carcinoma: the impact of rigid criteria to distinguish non-invasive f. *Mod Pathol* **30**:810–825.
7. Alves V, Kakudo K, LiVolsi V, Lloyd R, Nikiforov Y, Nosé V, Papotti M, Thompson L 2018 Noninvasive follicular thyroid neoplasm with papillary-like nuclear features (NIFTP): achieving better agreement by refining diagnostic criteria. *Clinics* **73**:1–4.
8. Kakudo K, El-Naggar AK, Hodak SP, Khanafshar E, Nikiforov YE, Nosé V, Thompson LDR 2018 Noninvasive follicular thyroid neoplasm with papillary-like nuclear features (NIFTP) in thyroid tumor classification. *Pathol Int* **68**:327–333.
9. Chen AY, Jemal A, Ward EM 2009 Increasing incidence of differentiated thyroid cancer in the United States, 1988–2005. *Cancer* **115**:3801–3807.
10. Jung CK, Little MP, Lubin JH, Brenner A V, Wells SA, Sigurdson AJ, Nikiforov YE 2014 The increase in thyroid cancer incidence during the last four decades is accompanied by a high frequency of BRAF mutations and a sharp increase in RAS mutations. *J Clin Endocrinol Metab* **99**:E276–E285.
11. Hirokawa M, Carney JA, Goellner JR, DeLellis RA, Hefless 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.
12. Lloyd RV, Erickson LA, Casey MB, Lam KY, Lohse CM, Asa SL, Chan JKC, DeLellis RA, Harach HR, Kakudo K, LiVolsi VA, Rosai J, Sebo TJ, Sobrinho-Simoes M, Wenig BM, Lae ME 2004 Observer variation in the diagnosis of follicular variant of papillary thyroid carcinoma. *Am J Surg Pathol* **28**:1336–1340.
13. Elsheikh TM, Asa SL, Chan JKC, DeLellis RA, Heffess CS, LiVolsi VA, Wenig BM 2008 Interobserver and intraobserver variation among experts in the diagnosis of thyroid follicular lesions with borderline nuclear features of papillary carcinoma. *Am J Clin Pathol* **130**: 736–744.
14. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, STROBE initiative 2007 the strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *Lancet (London, England)* **370**: 1453–1457.
15. McHugh ML 2012 Interrater reliability: the kappa statistic. *Biochem Med* **22**:276–282.
16. Xu B, Serrette R, Tuttle RM, Alzumaili B, Ganly I, Katabi N, Tallini G, Ghossein R 2019 How many papillae in conventional papillary carcinoma? A clinical evidence-based pathology study of 235 unifocal encapsulated papillary thyroid carcinomas, with emphasis on the diagnosis of noninvasive follicular thyroid neoplasm with papillary-like nucle. *Thyroid* **29**:1792–1803.
17. Bychkov A, Hirokawa M, Jung CK, Liu Z, Zhu Y, Hong SW, Satoh S, Lai C-R, Huynh L, Kakudo K 2017 Low rate of noninvasive follicular thyroid neoplasm with papillary-like nuclear features in asian practice. *Thyroid* **27**:983–984.
18. Ullmann TM, Gray KD, Stefanova D, Limberg J, Buicko JL, Finnerty B, Zarnegar R, Fahey TJ, Beninato T 2019 The 2015 American Thyroid Association guidelines are associated with an increasing rate of hemithyroidectomy for thyroid cancer. *Surgery* **166**:349–355.

Address correspondence to:
Frederick Thurston Drake, MD, MPH
Department of Surgery
Section of Endocrine Surgery
Boston University
Boston, MA 02118
 USA

E-mail: thurston.drake@bmc.org