

Serveur Académique Lausannois **SERVAL** serval.unil.ch

Author Manuscript

Faculty of Biology and Medicine Publication

This paper has been peer-reviewed but does not include the final publisher proof-corrections or journal pagination.

Published in final edited form as:

Title: Impact of non-invasive follicular thyroid neoplasms with papillary-like nuclear features (NIFTP) on risk of malignancy in patients undergoing lobectomy/thyroidectomy for suspicious for malignancy or malignant fine-needle aspiration cytology findings: a systematic review and meta-analysis.

Authors: Bongiovanni M, Faquin W, Giovanella L, Durante C, Kopp P, Trimboli P

Journal: European journal of endocrinology

Year: 2019 Jul 1

DOI: 10.1530/EJE-19-0223

In the absence of a copyright statement, users should assume that standard copyright protection applies, unless the article contains an explicit statement to the contrary. In case of doubt, contact the journal publisher to verify the copyright status of an article.

**Impact of non-invasive follicular thyroid neoplasms with papillary-like nuclear features (NIFTP) on risk of malignancy in patients undergoing lobectomy/thyroidectomy for suspicious for malignancy or malignant fine-needle aspiration cytology findings:
a systematic review and meta-analysis.**

Massimo Bongiovanni, MD¹; William C. Faquin, MD²; Luca Giovanella, MD³; Cosimo Durante, MD⁴;
Peter Kopp, MD⁵; Pierpaolo Trimboli, MD³.

Authors' affiliations:

¹Service of Clinical Pathology, Lausanne University Hospital, Institute of Pathology, Lausanne, Switzerland.

²Department of Pathology, Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts, USA.

³Department of Nuclear Medicine and Thyroid Centre, Ente Ospedaliero Cantonale, Bellinzona, Switzerland.

⁴Dipartimento di Medicina Traslazionale e di Precisione, Azienda Ospedaliera Universitaria Policlinico Umberto I, Sapienza Università di Roma, Rome, Italy.

⁵Service of Endocrinology, Diabetology and Metabolism, CHUV, Lausanne University Hospital, Lausanne, Switzerland.

Corresponding author:

Massimo Bongiovanni, MD, Service of Clinical Pathology, Lausanne University Hospital, Institute of Pathology, Rue du Bugnon 25, CH-1011 Lausanne, Switzerland.

Tel.: +41 21 314 72 02;

Fax: +41 21 314 72 05;

E-mail: massimo.bongiovanni@chuv.ch

SHORT TITLE: NIFTP impact on ROM: a meta-analysis.

Abstract

Objective. The second version of The Bethesda System for Reporting Thyroid Cytopathology endorsed the introduction of non-invasive follicular thyroid neoplasms with papillary-like nuclear features (NIFTP) as a distinct entity with low malignant potential into clinical practice. Consequently, the risk of malignancy (ROM) of cytological diagnoses has changed, but the magnitude of the change remains uncertain. The present systematic review was undertaken to obtain more robust information about the true impact of NIFTP on the ROM among patients undergoing surgery following a fine-needle aspiration cytology (FNAC) diagnosis of suspicious for malignancy (Bethesda V) or malignant (Bethesda VI). As they are managed surgically, these two diagnostic categories are the primary entities that are clinically impacted by the advent of NIFTP.

Design. Systematic review and meta-analysis.

Methods. A comprehensive literature search of online databases was performed in November 2018. The search was conducted looking for data of histologically proven NIFTP with preoperative FNAC.

Results. One-hundred-fifty-seven articles were identified and nine were included in the study. Overall, there were 13,752 thyroidectomies with a cancer prevalence of 45.7%. When NIFTP was considered non-malignant, the pooled risk difference for ROM was 5.5%. Applying meta-analysis, the pooled prevalence of NIFTP among nodules with FNAC of Bethesda V or Bethesda VI was 14% and 3%, respectively.

Conclusion. This meta-analysis shows that the inclusion of NIFTP leads to a reduction in the ROM for the Bethesda V and Bethesda VI FNAC diagnostic categories by 14% and 3%, respectively. Clinicians should be aware of these data to avoid overtreatment.

Introduction

Non-invasive follicular thyroid neoplasm with papillary-like nuclear features (NIFTP), formerly known as the non-invasive encapsulated follicular variant of papillary thyroid carcinoma (non-invasive EFVPTC), represents a controversial entity and a popular topic (1, 2). Interobserver variability in the diagnosis of NIFTP, analogous to its EFVPTC counterpart, is expected to be high, even among expert thyroid pathologists. Microfollicular structures and, most importantly, borderline-atypical nuclear features characteristic of these lesions, are sources of diagnostic discrepancies (3, 4, 5). Due to its indolent behavior, NIFTP should be considered as a neoplasm with a low risk of relapse (6). Accordingly, these patients can be treated by lobectomy without a need for completion thyroidectomy and radioiodine therapy (6, 7, 8). NIFTP can only be diagnosed on surgical specimens, and even if its histological and cytological diagnostic criteria are refined continuously, NIFTP identification by fine-needle aspiration cytology (FNAC) is currently not feasible in a reliable way (9). In fact, as demonstrated in a recent meta-analysis of 15 studies including 915 NIFTP, these cases may be classified in all diagnostic categories (DC) of The Bethesda System for Reporting Thyroid Cytopathology (TBSRTC): 3% were diagnosed as Nondiagnostic/Unsatisfactory (ND/U), 10% as Benign (B), 30% as Atypia of Undetermined Significance or Follicular Lesion of Undetermined Significance (AUS/FLUS), 21% as Follicular Neoplasm or Suspicious for a Follicular Neoplasm (FN/SFN), 24% as suspicious for malignancy (SM), and 8% as malignant (M) (10). Importantly, only mild heterogeneity between the studies was identified (10). These findings highlight that FNAC will not provide a definite diagnosis of NIFTP, and that its presence pre-operatively may be considered only as a possibility in the differential diagnosis.

Since the introduction of NIFTP, the second edition of TBSRTC considered NIFTP as a non-malignant tumor; as a consequence, the risk of malignancy (ROM) associated with its six diagnostic categories has been revised leading to a variable reduction of the ROM among the categories (11). Particularly, the expected ROM associated with TBSRTC DC Bethesda V and VI has decreased from 50-75% and 97-99%, to 45-60% and 94-96%, respectively, based on a limited number of retrospective studies (11). These changes could have a significant impact on the surgical management for these patients, as overtreatment should be avoided. In fact, in both FNAC DC, surgical excision of the lesions is recommended with the options of either a lobectomy or a thyroidectomy.

In this context, the focus of this study was to obtain more robust information about the true impact of NIFTP on ROM of patients undergoing surgery due to Bethesda V or VI lesions. Accordingly, our meta-analysis aimed to estimate 1) the difference in ROM among all surgical patients when considering NIFTP as a malignant versus non-malignant entity; and 2) the prevalence of NIFTP among patients undergoing thyroid surgery following a FNAC report of Bethesda V or VI.

Material and Methods

Conduct of review

This present systematic review was conducted according to Prisma guidelines (Supplementary Figure 1).

Search strategy

A comprehensive literature search was conducted using the online databases of Pubmed/MEDLINE and Scopus. The search aimed to find original studies describing the presence of NIFTP among thyroidectomized patients in which there was a presurgical FNAC report according to TBRSTC. A combination of the terms “non-invasive”, “encapsulated”, “follicular variant”, “NIFTP”, “histopathology” and “thyroid cancer” was used. This approach would allow identification of a large number of studies that reported histologically proven NIFTP, and which contained information on the presurgical cytological assessment. A beginning date limit was not used. The search was updated until November 15, 2018, and no language restrictions were used. With an attempt to expand the search, references of the retrieved articles were also screened to identify additional studies.

Study selection

As the main inclusion criterion, only original articles reporting NIFTP cases diagnosed histologically in a consecutive series of patients undergoing thyroid surgery over a specific study period were included. Subsequently, studies in which preoperative FNAC classification according to TBSRTC was not used were excluded. In addition, articles with overlapping patient or nodule data were excluded, and case reports were not considered. Two researchers (MB, PT) independently reviewed titles and abstracts of the retrieved articles, applying the selection criteria; then, all authors independently reviewed the full-text of the remaining articles to determine their final inclusion.

Data extraction

For each included study, the following information was extracted independently by two investigators (MB, PT) in a piloted form: 1) study data (authors, year and journal of publication, country of origin); 2) study period; 3) number of thyroid surgery; 4) number of cancers among patients; 5) number of NIFTP diagnosed by histology; 6) number of preoperative FNA reports of Bethesda V or VI; 7) number of cancers among nodules with preoperative FNA reports of Bethesda V or VI. Data were cross-checked, and any discrepancies were discussed and mutually solved.

Study quality assessment

The risk of bias of included studies was assessed independently by two reviewers (MB, PT) through the Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) tool for the following aspects: patient selection; index test; reference standard; flow and timing. Risk of bias and concerns about applicability were rated as low, high and unclear risk.

Statistical analysis

A proportion meta-analysis was performed to obtain the pooled rate of histologically proven NIFTP among all patients who underwent surgery and within the subgroup of cases with a cancer diagnosis. Particularly, a lesion-based analysis was conducted. Data input: they were entered in two columns the number of histologically proven NIFTP and the total number of histologies for each study. A risk difference (RD) meta-analysis was conducted to evaluate the difference of ROM when considering NIFTP as a malignant or non-malignant entity. For statistical pooling of data,

the DerSimonian and Laird method (random-effects model) was used (12). Pooled data are presented with 95% confidence intervals (95% CI) and displayed using a forest plot. The I-square index was used to quantify the heterogeneity among the studies, and significant heterogeneity was defined as an I-square value > 50%. Egger's test was carried out to evaluate the possible presence of a publication bias. A funnel plot was carried out for any results and publication bias might be considered when smaller size studies had on average different results with respect to the larger ones. Statistical analyses were performed using the StatsDirect statistical software (StatsDirect Ltd; Altrincham, UK).

Results

Eligible articles

After exclusion of duplicate publications, the comprehensive computer literature search retrieved 157 articles. Of these, 136 were excluded according to the above criteria after review of titles and abstracts, and 21 full-text papers were evaluated. Finally, nine articles were included for the purposes of this study (Figure 1) (13, 14, 15, 16, 17, 18, 19, 20, 21). One study presented a multi-institutional experience; one of the series (from University of Pennsylvania, Philadelphia, PA) was excluded due to the risk of overlapping of data with other studies (21).

Qualitative analysis (systematic review)

The included studies were published by authors from the USA, Asia, or Europe (13, 14, 15, 16, 17, 18, 19, 20, 21). Overall, 13,752 thyroidectomies were described with a thyroid cancer prevalence of 45.7%. The study design of these articles always included a retrospective

examination of thyroid histological samples from consecutive series of thyroidectomies performed over a specific time period. Data on the final histological follow-up were clearly identified in texts, figures or tables of the studies. This allowed calculation of the ROM considering NIFTP as malignant and the ROM considering NIFTP as a non-malignant entity. In addition, the prevalence of NIFTP among those lesions with a preoperative FNAC diagnosis of Bethesda V or VI were clearly reported in eight of the selected studies (13, 14, 15, 16, 17, 18, 20, 21). Table 1 details the main characteristics and findings of the nine included studies. The overall distribution of NIFTP according to TBRSTC was similar to that observed in a recent meta-analysis with the following percentages in the various diagnostic categories: Bethesda I 3.4%, Bethesda II 11.2%, Bethesda III 30.9%, Bethesda IV 23.7%, Bethesda V 22.5%, and Bethesda VI 8.3% (10).

Quantitative analysis (meta-analysis)

The overall RD of ROM when considering NIFTP as a non-malignant entity was estimated pooling the histological results of all nine included studies (Table 1) (13, 14, 15, 16, 17, 18, 19, 20, 21). Among 13,752 patients undergoing thyroid surgery, 696 (5.1%) NIFTP were identified at histology. The pooled RD of ROM was of 5.5% (95%CI from 3.2 to 7.8) (Figure 2); then, when NIFTP was considered as a non-malignant entity the risk of malignancy decreased by 5.5%.

As illustrated in Table 1, data of FNAC diagnosis for the overall series of patients undergoing surgery was reported in eight articles (as the study by Lastra et al. (19) did not report data about Bethesda V and VI) and the overall number of histologically proven NIFTP with preoperative FNAC data was 557 (13, 14, 15, 16, 17, 18, 20, 21). In these articles, there were 1,208 nodules with a FNAC diagnosis of Bethesda V and another 3,121 with Bethesda VI. Among all 4,329 cases, there

were 185 NIFTP. At meta-analysis, the pooled prevalence of NIFTP among nodules with a cytological classification of Bethesda V was 14% (95%CI from 6 to 26) (Figure 3). Also, the pooled prevalence of NIFTP among nodules classified as Bethesda VI was 3% (95%CI from 1 to 6) (Figure 4).

Study quality assessment

The risk of bias of the studies is shown in Table 2. Overall, a high risk of bias was found regarding the standard reference (i.e. histology). This risk of bias is high because the evaluation of histological samples was done before the introduction of NIFTP. In fact, to diagnose NIFTP, one needs an analysis of the entire capsule and the totality of the nodule content in the histological sample, because the examination of the capsule is an essential prerequisite to exclude invasion and analysis of the entire lesion is necessary to exclude the presence of any papillary structures (22). Moreover, a high risk of bias was detected for the index test (i.e. FNAC). In fact, the design of all the included studies comprised a retrospective re-evaluation of cytological specimens of lesions evaluated by FNAC before the introduction of NIFTP as a separate entity. Low risk was observed in patient selection and flow and timing of the studies.

Discussion

NIFTP is the designation for a thyroid neoplasm that was recently introduced to replace the nomenclature for lesions formerly classified as non-invasive EFVPTC, an entity with an indolent and non-aggressive behavior. Following this reclassification, it has been recommended that clinical management consisting of lobectomy without completion thyroidectomy and without

radioiodine therapy is an adequate and optimal intervention to manage these patients since the recurrence rate is less than 1% one to two decades after surgery (6, 22). However, this recommendation assumes that we might be able to diagnose NIFTP preoperatively. Yet, a recent meta-analysis confirmed what was apparent from the onset, namely that FNAC diagnoses associated with NIFTP may be found in all DC of TBSRTC, i.e. from ND/U (Bethesda I) to M (Bethesda VI). Thus, the preoperative identification of NIFTP remains an unrealized goal (10).

Important for clinical practice, it would be useful to know to what extent the introduction of NIFTP affects the ROM within the different cytological categories, assuming that this neoplasm is classified as non-malignant. This issue was recently addressed in a meta-analysis by Vuong et al. which showed for the first time the impact of NIFTP on the ROM in all TBSRTC DC (23). In addition, there should be a particular focus on those patients for whom surgery is unequivocally recommended, i.e. patients with lesions classified by FNA in the Bethesda V and VI categories, and for whom the histologic diagnosis of NIFTP is associated with the risk of overtreatment. The present study was undertaken with a different aim, i.e. to provide more robust information on this specific subset of thyroid cases, and the possible implications for clinical practice.

First, our analysis demonstrates that the ROM of a series of consecutive patients undergoing thyroid surgery for all causes decreased by 5% when we considered NIFTP as a non-malignant tumor. Secondly, when we evaluated the FNAC cases with a diagnosis of Bethesda V and VI, considering NIFTP as non-malignant, we found that the ROM decreased by 14% and 3%, respectively. These data are in line with the study by Vuong et al. who showed a decrease of ROM in these two DC of 16% and 2.2%, respectively (23). The almost similar decrease in the ROM in the Bethesda V category (14.2%) was also confirmed recently by the study by Ventura et al. This

study was not included in our meta-analysis as published after the period of the search of the literature performed for the manuscript and did not included cases in the malignant category (24).

The second edition of TBSRTC, taking into account the impact of NIFTP, emphasizes that the diagnostic criteria used in the M DC should be more strictly applied, limiting this diagnosis to cases with clear-cut cytological features of malignancy. The presence of any papillary structures with a nuclear score of 3, three or more nuclear pseudo-inclusions, psammoma bodies, or the presence of high-grade features (mitosis, necrosis, solid/trabecular pattern) exclude a NIFTP diagnosis and are diagnostic of PTC (2, 25, 26, 27). Such an approach is expected to limit the false positive diagnosis due to NIFTP in the Bethesda VI. Furthermore, it would not significantly alter the high ROM for this DC, which was estimated in the present meta-analysis as 3%. However, to avoid any possible false negative cytological NIFTP cases in the Bethesda VI, one would expect more cases to be classified in the Bethesda V, thus making this category particularly affected by the introduction of NIFTP. The ROM of the Bethesda V was 14% in the present meta-analysis, but it is based mostly on retrospective studies. Cytopathologists are now more conscious about the NIFTP entity and the potential consequences for the patient's management. Therefore, many are applying stricter criteria for the cytological diagnosis of PTC. For these reasons, in our opinion, an increasing number of cases not fulfilling the criteria for the Bethesda VI (including NIFTP) will be moved into the Bethesda V, and this is likely to reduce the ROM in the Bethesda Bethesda V to an extent that is even higher than the 14% identified in the analysis presented here. Moreover, for the same reason as above, we will expect a further decrease in the ROM also in the indeterminate DC, (AUS/FLUS and FN/SFN) that is currently estimated in the meta-analysis by

Vuong et al. to be at 32% and 30%, respectively (23). Purely follicular-patterned lesions that have papillary-like nuclear features with few nuclear pseudo-inclusions will no longer be classified as Bethesda VI. These assumptions are consistent with the findings of Ohori et al. who showed that the introduction of an institutional quality assurance policy and applying strict criteria to their FNAC diagnosis for PTC (i.e. nuclear grooves and ≥ 3 pseudonuclear inclusions), led to a reduction in the number of Bethesda VI readings on cytology - hence, avoiding false positive diagnosis – in lesions that were ultimately diagnosed as NIFTP by histology (25). Consequently, they were able to maintain a high ROM rate in their Bethesda VI (25). Endocrine surgeons should be aware that the decrease of 14% in the ROM for the Bethesda V is probably an underestimate. For this reason, the management should be adapted and be more conservative in selected low risk cases, e.g. in lesions with an absence of worrisome ultrasound features or favorable molecular signatures. In this situation, a lobectomy may be the optimal first-line intervention in order to avoid overtreatment. Of note, Mito et al. showed that adding an explanatory note to the Bethesda V, i.e. that NIFTP is within the differential diagnosis, reduces the number of total thyroidectomies, and hence improved clinical management (27).

Certain limitations and strengths of the present study must be mentioned. First, almost all publications included in this review were retrospective studies and the authors reviewed histological and cytological samples of NIFTP specifically for their study. Of course, this study design could introduce a bias. Ideally, one would evaluate the pooled results of prospective studies in which patients are managed according to the FNAC diagnosis. However, the studies included in this analysis presented data of patients managed before the introduction of NIFTP as a separate entity. Moreover, the difference in clinical practice between some Asian centers

(active surveillance for low-risk papillary thyroid carcinomas and watchful waiting for indeterminate thyroid cytologies) and the still prevalent Western approach (surgery for all Bethesda V and VI cytological diagnosis), and the consequently lower prevalence of NIFTP in Asian countries could have contributed to high heterogeneity in the present meta-analysis. Second, the histologic diagnosis of NIFTP could be compromised in samples collected before its introduction (i.e. retrospective studies), because pathologists may not have considered histologic issues related to the central parenchymal portion of the tumors, in addition to reviewing the entire capsule to exclude invasion. Third, since a subset of NIFTP can be cytologically classified in all cytological DC (10), one can conclude that there was a selection bias in the included studies since patients with cytological findings in the Bethesda I or II categories are typically not being managed surgically. On the other hand, this bias should be absent in Bethesda V or VI lesions, which are the focus of our present study, since patients with the latter lesions are typically undergoing surgical resection.

In conclusion, the present meta-analysis reveals that 5% of the surgically removed nodules are classified as NIFTP on histology, resulting in an overall reduction of the ROM. More importantly, the prevalence of histologically proven NIFTP among lesions with FNAC of Bethesda V and VI is 14% and 3%, respectively. Treating clinicians should be aware of these changes to the ROM, which, after considering other characteristics, may help individualizing patient management and avoiding overtreatment.

Declaration of interest

The authors declare that there is no conflict of interest that could be perceived as prejudicing the impartiality of the research reported.

Funding

This research did not receive any specific grant from any funding agency in the public, commercial or not-for-profit sector.

Author Contributions

Pierpaolo Trimboli, Luca Giovannella and Massimo Bongiovanni: conceptualization, investigation, original draft and writing. Pierpaolo Trimboli and Cosimo Durante: data curation, formal analysis, methodology. William C. Faquin and Peter Kopp: supervision and editing. All authors have reviewed the manuscript and agreed to their individual contributions prior to submission.

References

1. Baloch ZW, Seethala RR, Faquin WC, Papotti MG, Basolo F, Fadda G, Randolph GW, Hodak SP, Nikiforov YE & Mandel SJ. Noninvasive follicular thyroid neoplasm with papillary-like nuclear features (NIFTP): A changing paradigm in thyroid surgical pathology and implications for thyroid cytopathology. *Cancer Cytopathol* 2016 **124** 616-620.
2. Rossi ED & Faquin WC. NIFTP revised: Chronicle of a change foretold. *Cancer Cytopathol* 2018 **126** 897-901.
3. Lloyd RV, Erickson LA, Casey MB, Lam KY, Lohse CM, Asa SL, Chan JK, DeLellis RA, Harach HR, Kakudo K et al. Observer variation in the diagnosis of follicular variant of papillary thyroid carcinoma. *Am J Surg Pathol* 2004 **28** 1336-1340.
4. Hirokawa M, Carney JA, Goellner JR, DeLellis RA, Heffess CS, Katoh R, Tsujimoto M & Kakudo K. Observer variation of encapsulated follicular lesions of the thyroid gland. *Am J Surg Pathol* 2002 **26** 1508-1514.
5. Thompson LDR, Poller DN, Kakudo K, Burchette R, Nikiforov YE & Seethala RR. An International Interobserver Variability Reporting of the Nuclear Scoring Criteria to Diagnose Noninvasive Follicular Thyroid Neoplasm with Papillary-Like Nuclear Features: a Validation Study. *Endocr Pathol* 2018 **29** 242-249.
6. Nikiforov YE, Seethala RR, Tallini G, Baloch ZW, Basolo F, Thompson LD, Barletta JA, Wenig BM, Al Ghuzlan A, Kakudo K et al. Nomenclature Revision for Encapsulated Follicular Variant of Papillary Thyroid Carcinoma: A Paradigm Shift to Reduce Overtreatment of Indolent Tumors. *JAMA Oncol* 2016 **2** 1023-1029.
7. Rivera M, Ricarte-Filho J, Tuttle RM, Ganly I, Shaha A, Knauf J, Fagin J & Ghossein R. Molecular, morphologic, and outcome analysis of thyroid carcinomas according to degree of extrathyroid extension. *Thyroid* 2010 **20** 1085-1093.
8. Baloch ZW, Shafique K, Flannagan M & Livolsi VA. Encapsulated classic and follicular variants of papillary thyroid carcinoma: comparative clinicopathologic study. *Endocr Pract* 2010 **16** 952-959.
9. Lloyd RV, Asa SL, LiVolsi VA, Sadow PM, Tischler AS, Ghossein RA, Tuttle RM & Nikiforov YE. The evolving diagnosis of noninvasive follicular thyroid neoplasm with papillary-like nuclear features (NIFTP). *Hum Pathol* 2018 **74** 1-4.
10. Bongiovanni M, Giovanella L, Romanelli F & Trimboli P. Cytological diagnoses associated with non-invasive follicular thyroid neoplasms with papillary-like nuclear features, (NIFTP) according to the Bethesda System for Reporting Thyroid Cytopathology: a systematic review and meta-analysis. *Thyroid* 2018.
11. Ali SZ & Cibas ES. *The Bethesda System for Reporting Thyroid Cytopathology. Definitions, Criteria, and Explanatory Notes*. New York: Wiley, 2017.
12. DerSimonian R & Laird N. Meta-analysis in clinical trials revisited. *Contemp Clin Trials* 2015 **45** 139-145.
13. Strickland KC, Howitt BE, Marqusee E, Alexander EK, Cibas ES, Krane JF & Barletta JA. The Impact of Noninvasive Follicular Variant of Papillary Thyroid Carcinoma on Rates of Malignancy for Fine-Needle Aspiration Diagnostic Categories. *Thyroid* 2015 **25** 987-992.

14. Canberk S, Gunes P, Onenerk M, Erkan M, Kilinc E, Kocak Gursan N & Kilicoglu GZ. New Concept of the Encapsulated Follicular Variant of Papillary Thyroid Carcinoma and Its Impact on the Bethesda System for Reporting Thyroid Cytopathology: A Single-Institute Experience. *Acta Cytol* 2016 **60** 198-204.
15. Faquin WC, Wong LQ, Afrogheh AH, Ali SZ, Bishop JA, Bongiovanni M, Puzstaszeri MP, VandenBussche CJ, Gourmaud J, Vaickus LJ et al. Impact of reclassifying noninvasive follicular variant of papillary thyroid carcinoma on the risk of malignancy in The Bethesda System for Reporting Thyroid Cytopathology. *Cancer Cytopathol* 2016 **124** 181-187.
16. Layfield LJ, Baloch ZW, Esebua M, Kannuswamy R & Schmidt RL. Impact of the Reclassification of the Non-Invasive Follicular Variant of Papillary Carcinoma as Benign on the Malignancy Risk of the Bethesda System for Reporting Thyroid Cytopathology: A Meta-Analysis Study. *Acta Cytol* 2017 **61** 187-193.
17. Bychkov A, Keelawat S, Agarwal S, Jain D, Jung CK, Hong S, Lai CR, Satoh S & Kakudo K. Impact of non-invasive follicular thyroid neoplasm with papillary-like nuclear features on the Bethesda system for reporting thyroid cytopathology: a multi-institutional study in five Asian countries. *Pathology* 2018 **50** 411-417.
18. Kim M, Kim JE, Kim HJ, Chung YR, Kwak Y & Park SY. Cytologic Diagnosis of Noninvasive Follicular Thyroid Neoplasm with Papillary-like Nuclear Features and Its Impact on the Risk of Malignancy in the Bethesda System for Reporting Thyroid Cytopathology: An Institutional Experience. *J Pathol Transl Med* 2018 **52** 171-178.
19. Lastra RR, Birdsong G, Hwang DH, Jorda M, Kerr DA, McGrath C, Odronic S, Rao R, VanderLaan PA, Walker JW et al. Preoperative cytologic interpretation of noninvasive follicular thyroid neoplasm with papillary-like nuclear features: a 1-year multi-institutional experience. *Journal of the American Society of Cytopathology* 2018 **7** 79-85.
20. Lindeman BM, Nehs MA, Angell TE, Alexander EK, Gawande AA, Moore FD, Jr., Doherty GM & Cho NL. Effect of Noninvasive Follicular Thyroid Neoplasm with Papillary-Like Nuclear Features (NIFTP) on Malignancy Rates in Thyroid Nodules: How to Counsel Patients on Extent of Surgery. *Ann Surg Oncol* 2018.
21. Zhou H, Baloch ZW, Nayar R, Bizzarro T, Fadda G, Adhikari-Guragain D, Hatem J, Larocca LM, Samolczyk J, Slade J et al. Noninvasive follicular thyroid neoplasm with papillary-like nuclear features (NIFTP): Implications for the risk of malignancy (ROM) in the Bethesda System for Reporting Thyroid Cytopathology (TBSRTC). *Cancer Cytopathol* 2018 **126** 20-26.
22. Nikiforov YE, Baloch ZW, Hodak SP, Giordano TJ, Lloyd RV, Seethala RR & Wenig BM. Change in Diagnostic Criteria for Noninvasive Follicular Thyroid Neoplasm With Papillarylike Nuclear Features. *JAMA Oncol* 2018 **4** 1125-1126.
23. Vuong HG, Tran TTK, Bychkov A, Jung CK, Nakazawa T, Kakudo K, Katoh R & Kondo T. CLINICAL IMPACT OF NON-INVASIVE FOLLICULAR THYROID NEOPLASM WITH PAPILLARY-LIKE NUCLEAR FEATURES ON THE RISK OF MALIGNANCY IN THE BETHESDA SYSTEM FOR REPORTING THYROID CYTOPATHOLOGY: A META-ANALYSIS OF 14,153 RESECTED THYROID NODULES. *Endocr Pract* 2019 **5** 491-502.
24. Ventura M, Melo M, Fernandes G & Carrilho F. RISK OF MALIGNANCY IN THYROID CYTOLOGY: THE IMPACT OF THE RECLASSIFICATION OF NONINVASIVE FOLLICULAR

- THYROID NEOPLASM WITH PAPILLARY-LIKE NUCLEAR FEATURES (NIFTP). *Endocr Pract.* 2019 Mar 13. doi: 10.4158/EP-2018-0423. [Epub ahead of print].
25. Ohori NP, Wolfe J, Carty SE, Yip L, LeBeau SO, Berg AN, Schoedel KE, Nikiforov YE & Seethala RR. The influence of the noninvasive follicular thyroid neoplasm with papillary-like nuclear features (NIFTP) resection diagnosis on the false-positive thyroid cytology rate relates to quality assurance thresholds and the application of NIFTP criteria. *Cancer Cytopathol* 2017 **125** 692-700.
 26. Maletta F, Massa F, Torregrossa L, Duregon E, Casadei GP, Basolo F, Tallini G, Volante M, Nikiforov YE & Papotti M. Cytological features of "noninvasive follicular thyroid neoplasm with papillary-like nuclear features" and their correlation with tumor histology. *Hum Pathol* 2016 **54** 134-142.
 27. Mito JK, Alexander EK, Angell TE, Barletta JA, Nehs MA, Cibas ES & Krane JF. A modified reporting approach for thyroid FNA in the NIFTP era: A 1-year institutional experience. *Cancer Cytopathol* 2017 **125** 854-864.

Figure legends

Figure 1. Flow-chart of search of studies. Full explanation of the terms used for the search is reported in the text.

Figure 2. A. Pooled risk difference (95% CI) of ROM in all 13,752 thyroidectomies when considering NIFTP as non-malignant entity rather than malignant. The vertical line at point zero of the x axis represents the absence of difference. The result shows that if NIFTP is considered as a non-malignant tumor, the ROM is reduced by 5.5% (95% CI from 3 to 7). I² (inconsistency) was 78.6% (95% CI = 54.6% to 87.2%). B. As shown in funnel plot, a publication bias might be considered. Egger test: P = 0.067.

Figure 3. A. Pooled prevalence (95% CI) of NIFTP cases among those 1,208 nodules with a preoperative FNAC report of Bethesda V. The results show that NIFTP represents 14% (95% CI from 6 to 26) of all cases. I² (inconsistency) was 95.8% (95% CI = 94.2% to 96.8%). B. As shown in funnel plot, a publication bias might be considered. Egger test: P = 0.004.

Figure 4. A. Pooled prevalence (95% CI) of NIFTP cases among 3,121 nodules with a preoperative FNAC report of Bethesda VI. The results show that NIFTP represents 3% (95% CI from 1 to 6) of all cases. I² (inconsistency) was 92.7% (95% CI = 88.6% to 94.9%). B. As shown in funnel plot, a publication bias might be considered. Egger test: P = 0.004.

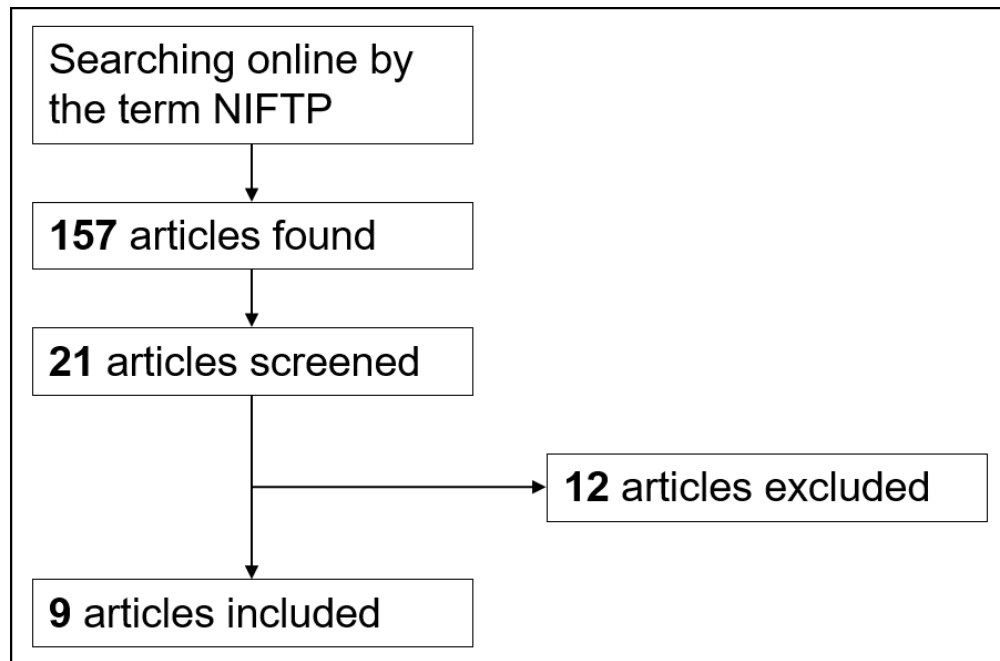


Figure 1. Flow-chart of search of studies. Full explanation of the terms used for the search is reported in the text.

152x100mm (150 x 150 DPI)

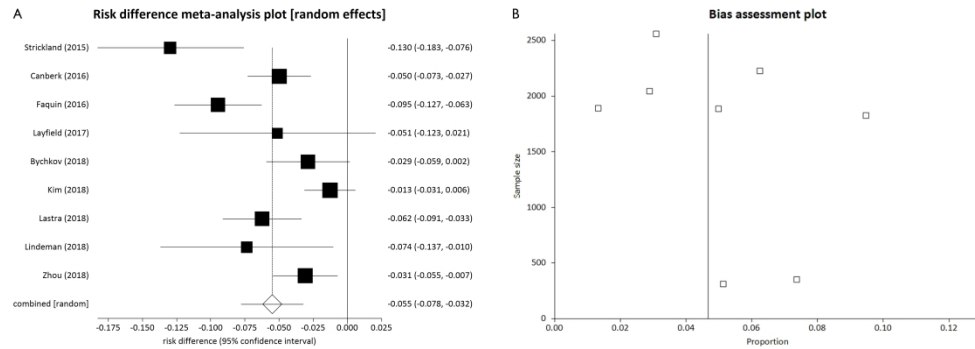


Figure 2. A. Pooled risk difference (95% CI) of ROM in all 13,752 thyroidectomies when considering NIFTP as non-malignant entity rather than malignant. The vertical line at point zero of the x axis represents the absence of difference. The result shows that if NIFTP is considered as a non-malignant tumor, the ROM is reduced by 5.5% (95% CI from 3 to 7). I² (inconsistency) was 78.6% (95% CI = 54.6% to 87.2%). B. As shown in funnel plot, a publication bias might be considered. Egger test: P = 0.067.

1196x448mm (120 x 120 DPI)

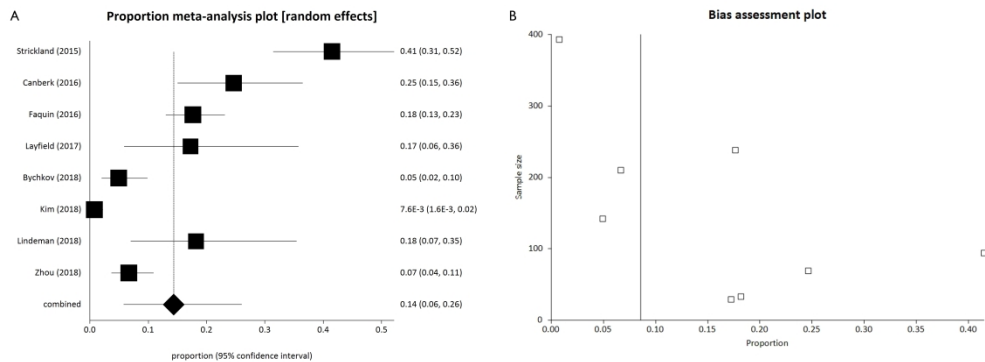


Figure 3. A. Pooled prevalence (95% CI) of NIFTP cases among those 1,208 nodules with a preoperative FNAC report of Bethesda V. The results show that NIFTP represents 14% (95% CI from 6 to 26) of all cases. I² (inconsistency) was 95.8% (95% CI = 94.2% to 96.8%). B. As shown in funnel plot, a publication bias might be considered. Egger test: P = 0.004.

1196x448mm (120 x 120 DPI)

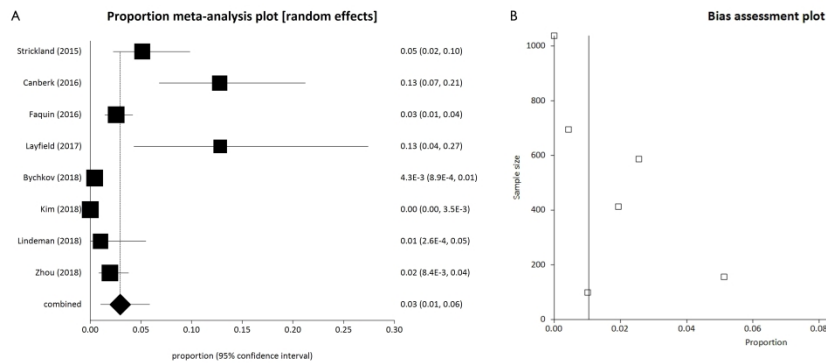


Figure 4. A. Pooled prevalence (95% CI) of NIFTP cases among 3,121 nodules with a preoperative FNAC report of Bethesda VI. The results show that NIFTP represents 3% (95% CI from 1 to 6) of all cases. I² (inconsistency) was 92.7% (95% CI = 88.6% to 94.9%). B. As shown in funnel plot, a publication bias might be considered. Egger test: P = 0.004.

1196x448mm (120 x 120 DPI)

Table 1. Main features of the included studies.

First author (year)	Cases reviewed	NIFTP found	ROM considering	ROM considering	Bethesda V	Bethesda VI
			NIFTP as a cancer (%)	NIFTP as non- malignant (%)		
Strickland (2015)	655	85	52.8	39.8	94	156
Canberk (2016)	1886	94	18.1	13.2	69	94
Faquin (2016)	1826	173	48.0	38.5	238	587
Layfield (2017)	312	16	32.6	27.5	29	39
Bychkov (2018)	2,044	59	54.3	51.4	142	695
Kim (2018)	1,891	25	91.2	89.9	393	1038
Lastra (2018)	2,226	139	45.0	38.8	NA	NA
Lindeman (2018)	353	26	28.0	20.6	33	99
Zhou (2018)	2559	79	26.4	23.3	210	413
Total	13,752	696	45.7	40.6	1,208	3,121

Legend – Cases reviewed represent all thyroid histological cases reviewed retrospectively in the studies. Bethesda V and Bethesda VI column report the overall cases with these FNA among all patients operated upon.

Table 2. Quality assessment of the studies according to QUADAS-2.

First author	Risk of bias				Feasibility		
	Patient selection	Index test	Reference standard	Flow and timing	Patient selection	Index test	Reference standard
Strickland (2015)	L	H	H	L	L	U	L
Canberk (2016)	L	H	H	L	L	U	L
Faquin (2016)	L	H	H	L	L	U	L
Layfield (2017)	L	H	H	L	L	U	L
Bychkov (2018)	L	H	H	L	L	U	L
Kim (2018)	L	H	H	L	L	U	L
Lastra (2018)	L	H	H	L	L	U	L
Lindeman (2018)	L	H	H	L	L	U	L
Zhou (2018)	L	H	H	L	L	U	L

Legend – The risk of bias was defined as low (L), high (H), or unclear (U).

Supplementary Table S1 : PRISMA Checklist.

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	1
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	2
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	3-4
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	NA
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	5
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	5
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	5
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	5
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	6
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	7
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	7
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	7
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	7
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	7

Supplementary Table S1 : PRISMA Checklist.

Page 1 of 2

Section/topic	#	Checklist item	Reported on page #
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	7
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	NA
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	8
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	8
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	9
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	8
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	8
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	8
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	NA
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	10
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	12
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	12
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	1

From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(7): e1000097. doi:10.1371/journal.pmed1000097

For more information, visit: www.prisma-statement.org.

Page 2 of 2