Endocrine (2019) 63:430–438
https://doi.org/10.1007/s12020-018-1825-8

META-ANALYSIS

Italian consensus for the classification and reporting of thyroid cytology: the risk of malignancy between indeterminate lesions at low or high risk. A systematic review and meta-analysis

Pierpaolo Trimboli[1] · Anna Crescenzi[2] · Marco Castellana[3] · Francesco Giorgino[3] · Luca Giovanella[1] · Massimo Bongiovanni[4]

Received: 20 August 2018 / Accepted: 25 November 2018 / Published online: 5 December 2018

Abstract

Background Italian consensus for the classification and reporting of thyroid cytology has proposed to discriminate the cancer prevalence of high (Tir 3B) vs. low (Tir 3A) risk indeterminate nodules. To obtain more robust evidence on this topic, we performed a meta-analysis of the Odds Ratio (OR) of malignancy of Tir 3B vs. Tir 3A nodules.

Methods A comprehensive literature exploration of online databases was conducted until May 2018. Original articles reporting histology of nodules cytologically classified as Tir 3A and Tir 3B were eligible. Pooled cancer prevalence in Tir 3A and Tir 3B, and OR of Tir 3B vs. Tir 3A were calculated.

Results The search revealed 95 articles, and 10 were included for the meta-analysis. Overall, 1168 indeterminate lesions were reported (441 Tir 3A and 727 Tir 3B), of which 391 were cancers. The pooled cancer prevalence was 17% in Tir 3A and 47% in Tir 3B. The OR of Tir 3B vs. Tir 3A was 4.24 (95% CI 2.75 to 6.53) with mild heterogeneity and without publication bias. When we considered non-invasive follicular thyroid neoplasms with papillary-like nuclear features (NIFTP) as non-malignant, cancer rate was lower, and OR of Tir 3B against Tir 3A was 2.93 (95% CI 1.60 to 5.37), with no heterogeneity but with publication bias.

Conclusions The Italian system for thyroid cytology is reliable to assess indeterminate lesions at low and high risk, being Tir 3B associated with a cancer risk significantly higher than Tir 3A, also when considering NIFTP as non-malignant entity.

Keywords Fine-needle aspiration (FNA) · Indeterminate nodules · Thyroid · Carcinoma · Risk of malignancy

Introduction

Cytology by fine-needle aspiration (FNA) under ultrasonographic (US) guidance (US-FNA) is recognized as a main tool for the evaluation of thyroid nodules [1, 2]. Thyroid FNA is very useful as a diagnostic test in benign and malignant nodules (macrofollicular/colloidal/inflammatory lesions and papillary thyroid carcinomas). Its main limitation is that indeterminate reports, usually represented by follicular lesions, can occur in up to 20–25% of various institutional series. Because only one in four of these nodules is expected to be a cancer, to preoperatively identify those nodules with lower risk represents a major challenge in clinical practice [3].

In the last decades, several international societies separated out two categories of indeterminate cytology to stratify such aspirates according to their probability of malignancy [4]. The British Thyroid Association (BTA)

Supplementary material The online version of this article (https://doi.org/10.1007/s12020-018-1825-8) contains supplementary material, which is available to authorized users.

✉ Pierpaolo Trimboli
pierpaolo.trimboli@eoc.ch

[1] Department of Nuclear Medicine and Thyroid Centre, Oncology Institute of Southern Switzerland, Bellinzona, Switzerland

[2] Section of Pathology, University Hospital Campus Bio Medico, Rome, Italy

[3] Department of Emergency and Organ Transplantation, Section of Internal Medicine, Endocrinology, Andrology and Metabolic Diseases, University of Bari Aldo Moro, Bari, Italy

[4] Service of Clinical Pathology, Institute of Pathology, Lausanne University Hospital, Lausanne, Switzerland

proposed the subclassification of indeterminate category into Thy 3a (low risk) and Thy 3f (high risk) [5], and The Bethesda System for Reporting Thyroid Cytopathology (TBSRTC) the subcategories atypia of undetermined significance/follicular lesion of undetermined significance (AUS/FLUS) at low risk, and follicular neoplasm/suspicious for a follicular neoplasm (FN/SFN) at high risk [6]. Later, the Italian consensus for the classification and reporting of thyroid cytology (ICCRTC) reported a five-class system with a subclassification of the indeterminate category Tir 3 into Tir 3A (low risk) and Tir 3B (high risk) [7]. Based on the previous experience of both BTA and TBSRTC, ICCRTC defined Tir 3A as increased cellularity with numerous microfollicular structures in a background of scant colloid, including here nodules with cytological or architectural changes (i.e., not atypia suspicious for papillary carcinoma), and Tir 3B as high cellularity in a monotonous microfollicular/trabecular arrangement with scant or absent colloid, also including in this subclass those cases with "mild/focal nuclear atypia" (cytological atypia). Compromised specimens (by smearing or staining artifacts) with cytological or architectural changes (i.e., not atypia suspicious for papillary carcinoma) are also included in the Tir 3A. The Tir 3B category includes specimens with high cellularity in a monotonous microfollicular/trabecular arrangement (architectural atypia), with scant or absent colloid. Even if the Italian consensus for reporting thyroid FNA cytology adopted a six-tiered class grouping, as a variance from the Bethesda Systems, the ICCRTC written in 2014 placed cases of indeterminate lesions with nuclear atypia (that formally cannot exclude the possibility of a papillary cancer) in the high risk indeterminate lesions. This choice of putting not only follicular pattern lesions (such as FN/SFN of TBSRTC), but also cytological alterations in the high-risk category was aimed to better stratify the malignancy risk in a crescent scale among categories. A recent meta-analysis found high difference between Tir 3A and Tir 3B in terms of prevalence of malignancy [8]. However, some potential limitations were present in that study and findings could be influenced by the inclusion of preliminary data published until that period [8].

To assess the true cancer prevalence in indeterminate lesions at low or high risk, all cases should be operated upon. In fact, when we refer to the risk of malignancy of these lesions, we are actually estimating the prevalence of cancers only among those cases that underwent surgical excision, due to managing of patients according to international recommendations, institutional guidelines, and other factors, such as clinical and surgical experience, as well as the patient's preference. Thus, the risk assigned to the category of indeterminate lesions by the above guidelines represents only an estimation reflecting available histological data following surgery [4–8].

Here we carried out an assessment of cancer rates recorded in Tir 3A and Tir 3B nodules. In particular, to skip the potential bias deriving from the different clinical management of indeterminate nodules in different institutions, we retrieved data to perform a meta-analysis on the Odds Ratio (OR) between the two subclasses. Accordingly, we conducted a systematic review of studies reporting the incidence of malignancy of thyroid nodules with prior FNA report of Tir 3A or Tir 3B, and a meta-analysis to calculate the pooled OR of Tir 3B vs. Tir 3A.

Material and methods

Conduct and registration of review

This present systematic review was conducted according to PRISMA guidelines [view Supplementary Table 1] and registered in PROSPERO (registration number CRD42018096937).

Search strategy

A comprehensive literature search in online databases of Google Scholar, Pubmed and Scopus was conducted by searching papers citing the Italian consensus for the classification and reporting of thyroid cytology (i.e., we searched by using the original title of the ICCRTC "Italian consensus for the classification and reporting of thyroid cytology" and then retrieved all papers with this article cited in the references list). This allowed us to retrieve studies reporting series of nodules that underwent FNA and were classified according to ICCRTC. A beginning date limit was not used, the search was updated until May 7 2018, and no language restriction was used. To try to expand our search, references of the included articles were also screened to identify additional studies.

Study selection

Articles found after our search were screened. Only original papers reporting complete data of nodules with indeterminate FNA report according to ICCRTC 2014 could be included in this systematic review. Exclusion criteria were: (a) articles not within the field of interest of this review; (b) articles that did not provide histological follow-up of the indeterminate nodules; (c) studies with overlapping patient or nodule data; (d) articles that did not provide clear study characteristics; (e) review articles, editorials, letters, or comments, case/series reports. After the initial selection of papers, we aimed to perform a second analysis. In fact, the cancer rate of thyroid FNA categories may be influenced by the recent re-classification, introduced into the 2017 WHO,

Table 1 Main characteristics of the included studies

First author [ref.]	Year	Journal	Country	Enrollment with respect to ICCRTC 2014 consensus	Indeterminate lesions found during the study period	Indeterminate lesions with histological follow-up (*n*)	Cancers at histology (*n*)	Cancer rate
Tartaglia [14]	2016	J Biol Regul Homeost Agents	Italy	Before	NA	52	14	26.9
Grani [15]	2017	Endocrine	Italy	Before	NA	49	19	38.8
Quaglino [16]	2017	Eur Thyroid J	Italy	Before	NA	96	37	38.5
Ulisse [17]	2017	Int J Endocrinol	Italy	Before	NA	50	15	30.0
Medas [18]	2017	Int J Surg	Italy	Before	NA	102	52	51.0
Straccia [19]	2017	Cytopathology	Italy	After	452	172	41	23.8
Rullo [20]	2018	J Endocrinol Invest	Italy	Before	NA	290	84	29.0
Trimboli [21]	2018	Endocr Pathol	Switzerland	After	63	51	9	17.6
Sparano [22]	2018	J Endocrinol Invest	Italy	After	655	273	101	37.0
Lauria [23]	2018	Eur J Endocrinology	Italy	After	244	33	19	57.6
Total						1168	391	33.5

All data refer to thyroid nodules with final histological examination. Enrollment of patients: studies designated as "before" collected a series of indeterminate lesions before the introduction of ICCRTC and re-classified that retrospectively; studies designated as "after" reported series of indeterminate nodules classified as Tir 3A or Tir 3B during the routine clinical practice and reviewed data for the publication

NA not available data

of the non-invasive follicular thyroid neoplasms with papillary-like nuclear features (NIFTP) as low malignant potential tumor [9]. Because older papers we retrieved could have not yet considered this change, NIFTP might be reported as malignant; in this second meta-analysis we considered NIFTP as a non-malignant entity. Articles in which data on NIFTP was not available were excluded for a secondary analysis to evaluate the impact of this entity on the odds of malignancy of Tir 3B over Tir 3A. Two researchers (PT, MC) independently reviewed titles and abstracts of the retrieved articles, applying the above 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 (PT, MC) in a piloted form: (1) study data (authors, year of publication and country of origin); (2) number of evaluated patients; (3) number of indeterminate lesions (Tir 3A, Tir 3B, total); (4) number of detected cancers and benignancies; (5) number of NIFTP detected in each category (when not available in the manuscript, two reviewers [PT or AC] contacted the corresponding authors to retrieve such information when possible). Data were cross-checked and any discrepancy discussed.

Study quality assessment

The risk of bias of included studies was assessed independently by two reviewers (PT, MC) 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 (L), high (H), or unclear (U).

Statistical analysis

The prevalence of cancers among the categories Tir 3A and Tir 3B was obtained from individual studies using the formula: number of cancers/number of thyroid nodules × 100. The ratio of malignancy rate in Tir 3A and Tir 3B was obtained in all papers, and then pooled. The OR between Tir 3B and Tir 3A was finally calculated. For statistical pooling of the data, DerSimonian and Laird method (random-effects model) was used [10]. In this model, pooled data represent weighted averages related to the sample size of the individual studies. Pooled data were presented with 95% confidence intervals (95% CI) and displayed using a forest plot. I-square index was used to quantify the heterogeneity among the studies as follows: <25%, no

heterogeneity; 25–50%, mild heterogeneity; 50–75%, moderate heterogeneity; >75%, high heterogeneity. Publication bias was assessed with Egger's test and funnel plot visually. In order to explore the heterogeneity of effect estimates, the event rate in Tir 3B vs. Tir 3A was displayed by L'Abbe plot and a meta-regression was performed [11, 12]. Statistical analyses were performed using the StatsDirect statistical software version 3.1.20 (StatsDirect Ltd; Altrincham, UK) and Review Manager (RevMan) version 5.3 (Computer program; Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2014), and ProMeta3 (Internovi).

Results

Eligible articles

The comprehensive computer literature search retrieved 95 articles. Review of titles and abstracts excluded 84 articles according to the above criteria, and 11 articles were initially included for the study (see supplementary Fig. 1) [13–23].

Qualitative analysis (systematic review)

All the included articles were published by Italian authors and have been critically reviewed by the authors before inclusion in this study. All included manuscripts were original papers, also that by Tartaglia et al. published as a "letter to the editor" [14]. In the manuscript by Grani et al. data of cancer rate in Tir 3A and Tir 3B were reported in Table and TP, FP, FN, TN are referred to Tir 3B vs. Tir 3A [15]. One paper was initially included because the authors calculated the cancer prevalence in Tir 3A and Tir 3B using histology as the gold standard; however, notably, these patients were managed according to the core-needle biopsy reports instead of FNA [13]. Some of these studies reported series of indeterminate nodules enrolled before the introduction of ICCRTC and re-classified retrospectively in Tir 3A or Tir 3B for the study [14–18, 20], while other studies included only nodules classified as Tir 3A or Tir 3B during the clinical practice [19, 21–23]. Data on the final histological follow-up were clearly identified in all these manuscripts, and the true percentage of malignancy could be calculated.

Quantitative analysis (meta-analysis)

Among the studies initially included in the systematic review ($n = 11$), one was excluded from the quantitative analysis due to the different management of patients during the clinical practice [13] and 10 were finally included for the meta-analysis. Table 1 details the characteristics and findings of these 10 studies and Table 2 describes their results.

Table 2 Histological results in Tir 3A and Tir 3B of the studies included in the meta-analysis

	Tir 3A		Tir 3B	
First author [ref.]	Cases	Cancers	Cases	Cancers
Tartaglia [14]	30	2 (6.7%)	22	12 (54.5%)
Grani [15]	23	6 (26.1%)	26	13 (50.0%)
		6 (*26.1%*)		13 (*50.0%*)
Quaglino [16]	44	9 (20.5%)	52	28 (53.8%)
Ulisse [17]	23	3 (13%)	27	12 (44.4%)
Medas [18]	19	4 (21.1%)	83	48 (57.8%)
Straccia [19]	63	10 (15.9%)	109	31 (28.4%)
		9 (*14.3%*)		27 (*24.8%*)
Rullo [20]	128	13 (10.2%)	162	71 (43.8%)
Trimboli [21]	40	3 (7.5%)	11	6 (54.5%)
		3 (*7.5%*)		5 (*45.5%*)
Sparano [22]	60	15 (25%)	213	86 (40.4%)
Lauria [23]	11	4 (36.4%)	22	15 (68.2%)
		4 (*36.4%*)		15 (*68.2%*)

All data refer to thyroid nodules with final histological examination. Cancers are reported as number (and percentage). Numbers in italics refer to data considering NIFTP as a non-malignant histological entity (view also Fig. 3)

The pooled rate of malignancy in all indeterminate lesions was 34%. The incidence of malignancy among the pooled series of indeterminate thyroid nodules of Tir 3A and Tir 3B was 17% and 47%, respectively ($p < 0.001$) (Table 3).

Since the cancer prevalence of Tir 3B was higher than Tir 3A, we could meta-analyze the OR of Tir 3B vs. Tir 3A. The pooled OR was 4.24 (95% CI 2.75 to 6.53) (Fig. 1), with mild heterogeneity ($I2 = 41.4$, 95% CI 0 to 70.5%). Figure 2 shows the L'Abbe plot of event rate in Tir 3B vs. Tir 3 A to explore the heterogeneity of effect estimated by the meta-analysis. Publication bias was absent (Egger test: 1.68, 95% CI = −0.94 to 4.3, $P = 0.177$) (Funnel plot is illustrated in supplemental Fig. 2). A secondary meta-analysis was performed separating articles into two group based on their study design; one group comprised 6 articles in which patients were enrolled before the introduction of ICCRTC and FNA samples were reviewed for the study aim [14–18, 20] and the other one included series in which ICCRTC was used during the routine clinical practice [19, 21–23]. As shown in Fig. 1, OR of Tir 3B vs. Tir 3A was significant high in both instances. To explore the heterogeneity found in the whole series, we performed a meta-regression considering the above subgroups (supplemental Fig. 3) and no statistical significance was found ($p = 0.132$).

Table 3 Histological prevalence of cancers among thyroid nodules with prior low or high risk indeterminate FNA report: meta-analysis of 1168 cases

	Cancer prevalence (95% CI)	Consistency–I2 (95% CI)	Publication bias–Egger test (95% CI)
All cases	34% (28 to 41)	78.6% (56.8 to 86.9)	1.68 (−2.56 to 5.91) $P = 0.3882$
Low risk (Tir 3A)	17% (12 to 22)	46.1% (0 to 72.5)	1.98 (0.25 to 3.70) $P = 0.0295$
High risk (Tir 3B)	47% (40 to 55)	67.5% (22.8 to 81.7)	2.17 (−0.43 to 4.76) $P = 0.0906$

Pooled cancer prevalence was calculated as random-effect model. Heterogeneity was defined as absent (I2 < 25%), mild (25–50%), moderate (50–75%), and high (>75%). Egger test significant identified presence of publication bias

In addition, we re-made the meta-analysis considering only those papers in which we had available data on NIFTP (see Table 2); particularly, in this meta-analysis we considered NIFTP as a benign histological outcome. The overall cancer rate was 24.5%, being 14% in Tir 3A and 33% in Tir 3B. The pooled OR of Tir 3B against Tir 3A was 2.93 (95% CI 1.60 to 5.37), with no heterogeneity and with publication bias (Egger test: 4.22, 95% CI = 0.81 to 7.64, $P = 0.0336$) (Fig. 3).

Finally, we evaluated the rate of nodules resected. As shown in Table 1, some studies did not report these data. Then, when we considered only available data from four articles [19, 21–23] and found 1414 indeterminate nodules of which 529 (37.4%) resected. The number of nodules resected among Tir 3A (174/875, 19.9%) was significantly lower than that observed in Tir 3B (355/539, 65.9%) ($p < 0.001$).

Study quality assessment

The risk of bias of the included studies is shown in Table 4. Overall, a low risk of bias was found: all consecutive patients diagnosed with indeterminate thyroid nodules in a specific period were included; FNA was obviously conducted and interpreted before histology; however, as a remarkable issue, in six articles FNA samples were reviewed and re-classified according to the ICCRTC only for the publication [14–18, 20]. Reference standard bias was rated as high since histology is commonly performed in knowledge of the results of the index test. Flow and timing bias was rated as low since thyroid cancer is a chronic condition.

Discussion

The present systematic review found a substantial number of articles reporting histological data of 1168 thyroid nodules with preoperative FNA diagnosis of Tir 3A or Tir 3B according to ICCRTC [7]. These data allowed us to perform a meta-analysis of cancer prevalence of Tir 3A and Tir 3B, of the OR of Tir 3B vs. Tir 3A, and further analyses of subgroups. The sample size and prevalence of cancer in both categories showed low-moderate heterogeneity, and this supports a sound finding. First, the present results confirm those observed in a previous meta-analysis on the preliminary published data, being the category of low-risk (Tir 3A) and high-risk (Tir 3B) indeterminate lesions associated with different cancer prevalence (17% and 47%, respectively) [8]. These findings are encouraging for all users of ICCRTC in their clinical practice. It is worth to underline that a meta-analysis on the reliability of TBSRTC did not find actual differences between the low-risk and high-risk categories, being the rate of malignancy 27% in AUS/FLUS and 31% in FN/SFN [24]. Another study on the diagnostic performance of pooled studies using TBSRTC showed a positive predictive value (PPV) for the AUS/FLUS diagnostic category of 15.9% and a PPV for the FN/SFN diagnostic category of 26.1% [25]. The higher difference in the malignancy risk between the two indeterminate subcategories in TBSRTC and Italian system (ie AUS/FLUS vs. FN/SFN and Tir 3A vs. Tir 3B, respectively) should be due to the different interpretation of nuclear atypia which are included into AUS/FLUS (low-risk) and Tir 3B (high-risk) category. The difference between the risk of malignancy of Tir 3B vs. Tir 3A in the current analysis using the ICCRTC are 30% compared to 10% in the meta-analysis of TBSRTC. Two recent meta-analysis on studies using the Bethesda system for reporting thyroid cytopathology evaluated the odds of malignancy of nuclear/cytologic atypia over the rate of malignancy of other types of atypia in aspirates with indeterminate cytology [26, 27]. Both studies found that the odds of malignancy were around 2.5-fold higher for aspirates with nuclear/cytologic atypia and suggested the need for standardizing the definition of nuclear atypia and for identifying it in the cytology reports. These data support the decision of the ICCRTC of incorporating nuclear/cytologic atypia into the high-risk (Tir 3B) category. Second, when we evaluated the OR of Tir 3B vs. Tir 3A, we found that the Tir 3B class has a risk of malignancy increased with respect to that of Tir 3A. On one hand, this result corroborates the above data, i.e., that Tir 3B

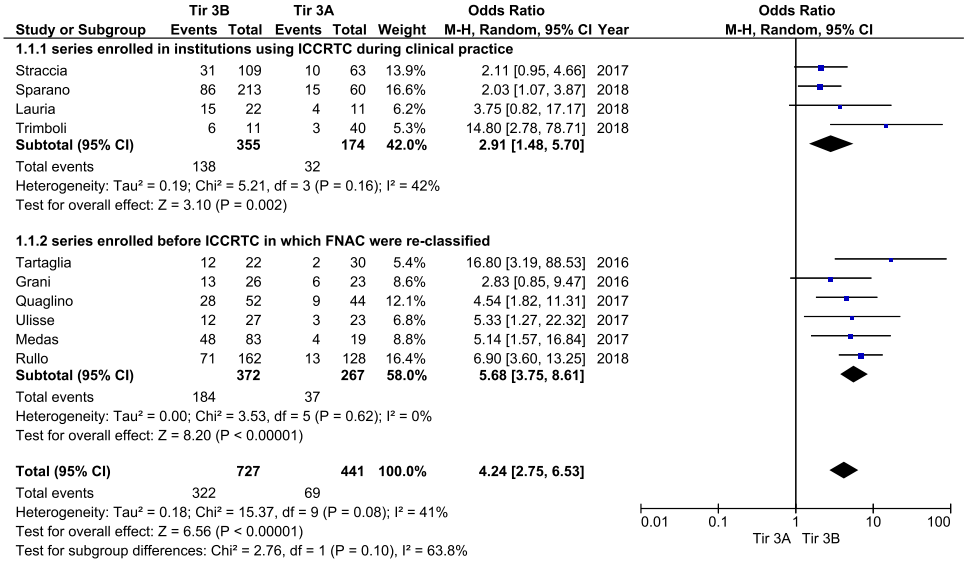

Fig. 1 Pooled OR of cancer (and 95% CI) in Tir 3B against Tir 3A. In the figure is reported the pooled OR of cancer (and 95% CI) in Tir 3B against Tir 3A also considering the subgroups of studies (1, studies reporting data of FNA samples reviewed according to the ICCRTC; 2, studies using ICCRTC during clinical practice)

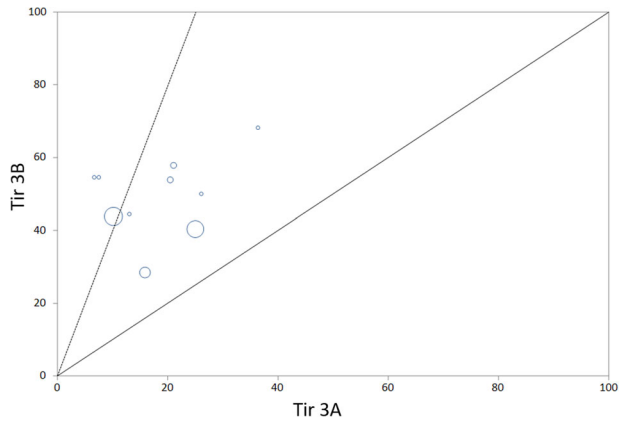

Fig. 2 L'Abbé plot of OR of Tir 3B vs. Tir 3A. Each circle represents an individual study and larger circles identify series with more participants. Solid diagonal line indicates the absence of difference of risk between Tir 3B and Tir 3A (OR = 1.0). The dotted line represents the pooled OR of all included studies

and Tir 3A have different cancer prevalence at histology; moreover, it demonstrates that the risk of a nodule classified as Tir 3B to be cancer is much higher than that of a nodule classified as Tir 3A regardless of the institutional management. Indeed, the evaluation of OR of Tir 3B vs. Tir 3A should delete the potential selection bias which affects the calculation of prevalence of cancer in the two classes. These data are clearly expressed by Fig. 2. Of relevance, the low intra-category heterogeneity observed for the malignancy rate for Tir 3A and Tir 3B, demonstrated a good attainment of the morphologic inclusion criteria among Italian cytopathologists. This means that splitting indeterminate lesions in two subclasses overcomes the previous difficulties in the identification of cytological samples with indeterminate features in which the risk of malignancy ranged from 5–15 to 30–50% and entailed a relevant clinical problem. Interestingly, this finding was confirmed when we analyzed separately studies according to their use of ICCRTC during clinical practice or not (Fig. 1). These data justify different management for the indeterminate cytological cases, i.e., repeat FNA vs. diagnostic surgery. Moreover, in selected cases with repeatedly Tir 3A diagnosis and no US and/or clinical suspicious findings, molecular test negativity could be a valid option to confirm a "wait and see" attitude and avoid surgery. As stated by the ICCRTC, clinical assessment is pivotal for surgical consultation while molecular test should be only an additional information [7]. We could better stratify Tir 3A cases using different qualifiers with associated different malignancy risk inside the Tir 3A category, ie discriminating architectural atypia (microfollicular structures), degenerative and regenerative changes, compromised specimens, and oncocytic features as it was suggested in the revised Bethesda system [28, 29]. The Italian system has already separated cases with cytological atypia into a different category (Tir 3B) [26, 27, 30]. All in all, the herein recorded data demonstrate a good performance of ICCRTC, which could be due to the different interpretation of atypia of papillary carcinoma; the latter are features of the high-risk subcategory (Tir 3B) of ICCRTC, while are included in the low-risk one of British (Thy 3a) and Bethesda (AUS/FLUS) systems. Although mild nuclear changes have been included in FN/SFN category in the last Edition of the TBSRTC, cellular atypia remains a criterion of AUS/FLUS, probably reducing the discriminating power of subcategories. We advise for further studies on this topic [28].

We conducted a secondary analysis to evaluate the weight of NIFTP on the malignancy rates of the two categories (Tir 3A and Tir 3B). For this purpose, NIFTP was considered non-malignant, and we excluded all those

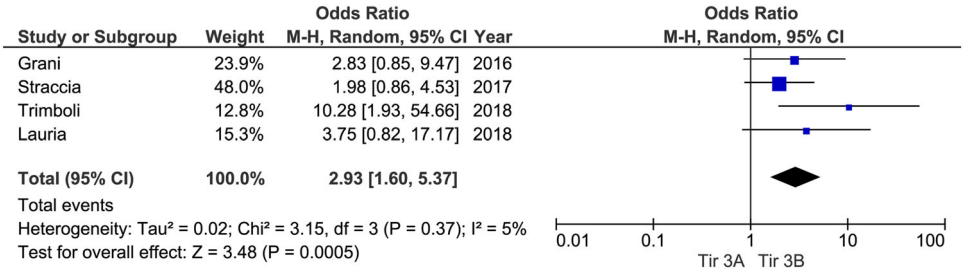

Fig. 3 Pooled OR of cancer (and 95% CI) in Tir 3B against Tir 3A considering NIFTP a non-malignant entity. Data on NIFTP were described in the published manuscript in one study and were known in other two studies. In the other cases the authors of this review contacted the authors of the original papers to obtain data on NIFTP

Table 4 Quality assessment of the studies according to QUADAS-2

Risk of bias					Feasibility		
First author [ref.]	Patients selection	Index test	Reference standard	Flow and timing	Patients selection	Index test	Reference standard
Tartaglia [14]	U	H	H	L	L	L	L
Grani [15]	U	H	H	L	L	L	L
Quaglino [16]	U	H	H	L	L	L	L
Ulisse [17]	U	H	H	L	L	L	L
Medas [18]	U	H	H	L	L	L	L
Straccia [19]	L	L	H	L	L	L	L
Rullo [20]	U	H	H	L	L	L	L
Trimboli [21]	L	L	H	L	L	L	L
Sparano [22]	L	L	H	L	L	L	L
Lauria [23]	L	L	H	L	L	L	L

L low risk of bias, *H* high risk of bias, *U* unclear risk of bias

studies in which the prevalence of NIFTP was not available/known. As expected, we found a decrease in the prevalence of malignancy in both categories (14% in Tir 3A and 33% in Tir 3B), and OR of Tir 3B against Tir 3A decreased to 2.93, reflecting a predominance of NIFTP in the Tir 3B category. This finding makes sense given that NIFTP would typically show a microfollicular proliferation and/or mild/moderate nuclear atypia on the cytology specimen, which would be classified into the Tir 3B category.

Limits and strengths of the present review have to be discussed. Generally, small sample-size studies reporting positive findings are more likely to report positive relationship and be published than those describing negative results. A selection bias might be present because these papers reported only patients who underwent surgery to obtain a definite histological diagnosis of the indeterminate lesions. However, the analysis of the ratio of cancer prevalence in Tir 3B and Tir 3A could delete this bias. Also, among the overall series of nodules included in the present meta-analysis, about two thirds were at high risk; particularly, if we consider only the four papers with series collected after the introduction of ICCRTC in which FNA were classified during clinical practice as Tir 3A and Tir 3B [19, 21–23] we found that in three studies of them the Tir 3B:Tir 3A ratio is higher (Tir 3B:Tir 3A ratio 2.0:1). This finding could reflect a more frequent surgical address in lesions at high risk than in the low-risk ones and could represent a potential bias affecting each study and the present one. In fact, here we found that about one third of indeterminate nodules were resected, being this rate significantly different between Tir 3A (19.9%) and Tir 3B (65.9%). We may speculate that the cancer rate recorded after surgery should be underestimated in Tir 3A while not in Tir 3B. We should discuss another potential bias of our first analysis; there, the cancer rate may be influenced by the presence of NIFTP among malignant cases [9, 19]. As a proof, when we performed a second sub-analysis of papers in which NIFTP data were available, we considered NIFTP as a non-malignant entity and we found a lower cancer rate in both low and high risk indeterminate lesions.

Relevant difference should be present between our previous meta-analysis [8] and the present one; (1) the present study includes a number of papers and cases significantly higher (1168 here vs. 423 there); (2) the present study includes also papers with series collected after the introduction of ICCRTC in which FNA were classified during clinical practice as Tir 3A and Tir 3B, such as a prospective selection of cases and not a retrospective review of Tir 3 to be re-classified in Tir A and Tir 3B (those patients were operated upon only due to a Tir 3 report); (3) the herein used statistical approach is significantly different because we aimed to calculate the OR between the two subclasses to skip the potential bias deriving from the different clinical management of indeterminate nodules in different institutions; (4) finally, here we performed a sub-meta-analysis of subgroup of studies considering the new entity of NIFTP as a non-malignant lesions, being this result a novelty.

In conclusion, this analysis confirms that the Italian system for classification of thyroid FNA can discriminate the indeterminate lesions at low risk of malignancy from that at high risk. Furthermore, Tir 3B is associated with a risk of malignancy significantly higher than Tir 3A, also when considering the new histological entity of NIFTP as a non-malignant entity. These results are of high interest for clinical practice.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

References

1. H. Gharib, E. Papini, R. Paschke et al. 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. Endocr. Pract. **16**(Suppl 1), 1–43 (2010)
2. B.R. Haugen, E.K. Alexander, K.C. Bible et al. 2015 American Thyroid Association Management Guidelines for Adult Patients with Thyroid Nodules and Differentiated Thyroid Cancer: The American Thyroid Association Guidelines Task Force on Thyroid Nodules and Differentiated Thyroid Cancer. Thyroid **26**, 1–133 (2016)
3. Z.W. Baloch, S. Fleisher, V.A. LiVolsi, P.K. Gupta, Diagnosis of "follicular neoplasm": a gray zone in thyroid fine-needle aspiration cytology. Diagn. Cytopathol. **26**, 41–44 (2002)
4. E.K. Alexander, G.C. Kennedy, Z.W. Baloch et al. Preoperative diagnosis of benign thyroid nodules with indeterminate cytology. N. Engl. J. Med. **367**, 705–715 (2012)
5. P. Perros, K. Boelaert, S. Colley et al. Guidelines for the management of thyroid cancer. Clin. Endocrinol. **81**(Suppl 1), 1–122 (2014)
6. E.S. Cibas, S.Z. Ali, The Bethesda system for reporting thyroid cytopathology. Thyroid **19**, 1159–1165 (2009)
7. F. Nardi, F. Basolo, A. Crescenzi et al. Italian consensus for the classification and reporting of thyroid cytology. J. Endocrinol. Invest. **37**, 593–599 (2014)
8. P. Trimboli, A. Crescenzi, L. Giovanella, Performance of Italian Consensus for the Classification and Reporting of Thyroid Cytology (ICCRTC) in discriminating indeterminate lesions at low and high risk of malignancy. A systematic review and meta-analysis. Endocrine **60**, 31–35 (2018)
9. Y.E. Nikiforov, R.A. Ghossein, K. Kakudo et al. in *Non-invasive follicular thyroid neoplasm with papillary-like nuclear features*. World Health Organization Classification of Tumours of Endocrine Organs. ed. by R.V. Lloyd, R.Y. Osamura, G. Kloppel, J. Rosai. (IARC, Lyon, France, 2017) pp. 78–80
10. R. DerSimonian, N. Laird, Meta-analysis in clinical trials. Control. Clin. Trials **7**, 177–188 (1986)
11. F. Song, Exploring heterogeneity in meta-analysis: Is the L'Abbé plot useful? J. Clin. Epidemiol. **52**, 725–730 (1999)
12. K.A. L'Abbé, A.S. Detsky, K. O'Rourke, Meta-analysis in clinical research. Ann. Intern. Med. **107**, 224–233 (1987)
13. P. Trimboli, L. Guidobaldi, S. Amendola et al. Galectin-3 and HBME-1 improve the accuracy of core biopsy in indeterminate thyroid nodules. Endocrine **52**, 39–45 (2016)
14. F. Tartaglia, A. Giuliani, L. Tromba et al. Fine needle aspiration cytology of 650 thyroid nodules operated for multinodular goiter: a cyto-histological correlation based on the new Italian cytological classification (SIAPEC 2014). J. Biol. Regul. Homeost. Agents **30**, 1187–1193 (2016)
15. G. Grani, L. Lamartina, V. Ascoli et al. Ultrasonography scoring systems can rule out malignancy in cytologically indeterminate thyroid nodules. Endocrine **57**, 256–261 (2017)
16. F. Quaglino, V. Marchese, E. Mazza, C. Gottero, R. Lemini, S. Taraglio, When is thyroidectomy the right choice? Comparison between fine-needle aspiration and final histology in a single institution experience. Eur. Thyroid J. **6**, 94–100 (2017)
17. S. Ulisse, D. Bosco, F. Nardi et al. Thyroid imaging reporting and data system score combined with the new italian classification for thyroid cytology improves the clinical management of indeterminate nodules. Int. J. Endocrinol. **2017**, 9692304 (2017)
18. F. Medas, E. Erdas, L. Gordini et al. Risk of malignancy in thyroid nodules classified as TIR-3A: what therapy? Int. J. Surg. **41**(Suppl 1), S60–S64 (2017)
19. P. Straccia, A. Santoro, E.D. Rossi et al. Incidence, malignancy rates of diagnoses and cyto-histological correlations in the new Italian Reporting System for Thyroid Cytology: an institutional experience. Cytopathology **28**, 503–508 (2017)
20. E. Rullo, G. Minelli, D. Bosco, F. Nardi, V. Ascoli, Evaluation of the Italian cytological subclassification of thyroid indeterminate nodules into TIR-3A and TIR-3B: a retrospective study of 290 cases with histological correlation from a single institution. J. Endocrinol. Invest. **41**, 531–538 (2018)
21. P. Trimboli, F. Fulciniti, E. Merlo, J. Barizzi, L. Mazzucchelli, L. Giovanella, Histologic outcome of indeterminate thyroid nodules classified at low or high risk. Endocr. Pathol. **29**, 75–79 (2018)
22. C. Sparano, G. Parenti, A. Cilotti et al. Clinical impact of the new SIAPEC-IAP classification on the indeterminate category of thyroid nodules. J. Endocrinol. Invest. (2018). https://doi.org/10.1007/s40618-018-0871-7
23. A. Lauria, E. Maddaloni, S.I. Briganti et al. Differences between ATA, AACE/ACE/AME and ACR TI-RADS ultrasound classifications performance in identifying cytological high-risk thyroid nodules. Eur. J. Endocrinol. **178**, 595–603 (2018)
24. P. Straccia, E.D. Rossi, T. Bizzarro et al. A meta-analytic review of the Bethesda system for reporting thyroid cytopathology: has the rate of malignancy in indeterminate lesions been underestimated? Cancer Cytopathol. **123**, 713–722 (2015)

25. M. Bongiovanni, A. Spitale, W.C. Faquin, L. Mazzucchelli, Z.W. Baloch, The Bethesda system for reporting thyroid cytopathology: a meta-analysis. Acta Cytol. **56**, 333–339 (2012)
26. S.H. Ahn, S.D. Kim, W.J. Jeong, Comparison of risk of malignancy in a subgroup with atypia of undetermined significance/follicular lesion of undetermined significance: a meta-analysis. Head Neck **39**, 1699–1710 (2017)
27. P. Valderrabano, L. Khazai, Z.J. Thompson, S.C. Sharpe, V.D. Tarasova, K.J. Otto, J.E. Hallanger-Johnson, J.T. Wadsworth, B. M. Wenig, C.H. Chung, B.A. Centeno, B. McIver, Cancer risk associated with nuclear atypia in cytologically indeterminate thyroid nodules: a systematic review and meta-analysis. Thyroid **28**, 210–219 (2018)
28. S.Z. Ali, E. Cibas. *The Bethesda System for Reporting Thyroid Cytopathology*, 2nd edn. (Springer, Cham, Switzerland, 2017)
29. M. Pusztaszeri, E.D. Rossi, M. Auger, Z. Baloch, J. Bishop, M. Bongiovanni, A. Chandra, B. Cochand-Priollet, G. Fadda, M. Hirokawa, S. Hong, K. Kakudo, J.F. Krane, R. Nayar, S. Parangi, F. Schmitt, W.C. Faquin, The Bethesda system for reporting thyroid cytopathology: proposed modifications and updates for the second edition from an international panel. Acta Cytol. **60**, 399–405 (2016)
30. C. Bellevicine, R. Sgariglia, I. Migliatico et al. Different qualifiers of AUS/FLUS thyroid FNA have distinct BRAF, RAS, RET/PTC, and PAX8/PPARg alterations. Cancer Cytopathol. **126**, 317–325 (2018)