META-ANALYSIS

The dilemma of ^{18}F-FDG PET/CT thyroid incidentaloma: what we should expect from FNA. A systematic review and meta-analysis

Lorenzo Scappaticcio[1] · Arnoldo Piccardo[2] · Giorgio Treglia[3,4,5,6] · David N. Poller[7,8] · Pierpaolo Trimboli[9,10]

Abstract

Purpose ^{18}F-FDG thyroid incidentaloma (TI) occurs in ~2% of PET/CT examinations with a cancer prevalence of up to 35–40%. Guidelines recommend fine-needle aspiration cytology (FNA) if a focal ^{18}F-FDG TI corresponds to a sonographic nodule >1 cm. The aim of this systematic review and meta-analysis was to provide evidence-based data on the diagnostic distribution of ^{18}F-FDG TIs in the six Bethesda systems for reporting thyroid cytopathology (BETHESDA) subcategories.
Methods Original studies reporting ^{18}F-FDG TIs and cytologically classified according to BETHESDA were included. Six separate meta-analyses were performed to obtain the pooled prevalence (95% confidence interval, 95% CI) of ^{18}F-FDG TIs in the six BETHESDA subcategories.
Results Fifteen studies were finally included. Nine studies were from Asian/Eastern and six from Western countries. FNA data according to BETHESDA was available in 2304 cases. The pooled prevalence of ^{18}F-FDG TIs according to BETHESDA was BETHESDA I 10% (6–14), BETHESDA II 45% (37–53), BETHESDA III 8% (3–13), BETHESDA IV 8% (5–12), BETHESDA V 6% (4–9), BETHESDA VI 19% (13–25). A significantly different prevalence was found in the BETHESDA IV between Asian/Eastern (2%) and Western (19%) studies.
Conclusion Two-thirds of focal ^{18}F-FDG TIs undergoing FNA have either malignant (BETHESDA VI) or benign (BETHESDA II) cytology while a minority will have indeterminate (BETHESDA III or IV) FNA results. Significant differences between Asian/Eastern and Western studies are also present in the prevalence of indeterminate FNA results.

Keywords Thyroid cancer · ^{18}F-FDG · PET/CT incidentaloma · FNA · Bethesda

These authors contributed equally: Lorenzo Scappaticcio, Arnoldo Piccardo

These authors jointly supervised this work: David N. Poller, Pierpaolo Trimboli

Supplementary information The online version contains supplementary material available at https://doi.org/10.1007/s12020-021-02683-4.

✉ Lorenzo Scappaticcio
lorenzo.scappaticcio@unicampania.it

1 Unit of Endocrinology and Metabolic Diseases, University of Campania 'L. Vanvitelli'', Naples, Italy

2 Department of Nuclear Medicine, Galliera Hospital, Genoa, Italy

3 Department of Nuclear Medicine and Molecular Imaging, Lausanne University Hospital, Lausanne, Switzerland

4 Faculty of Biology and Medicine, University of Lausanne, Lausanne, Switzerland

5 Clinic of Nuclear Medicine and Molecular Imaging, Imaging Institute of Southern Switzerland, Ente Ospedaliero Cantonale, Bellinzona, Switzerland

6 Health Technology Assessment Unit, Academic Education, Research and Innovation Area, Ente Ospedaliero Cantonale, Bellinzona, Switzerland

7 Departments of Cytology & Pathology, Queen Alexandra Hospital, Portsmouth, UK

8 UCL Cancer Institute, 72 Huntley St., Bloomsbury, London WC1E 6DD, UK

9 Clinic for Endocrinology and Diabetology, Lugano Regional Hospital, Ente Ospedaliero Cantonale, Lugano, Switzerland

10 Faculty of Biomedical Sciences, Università della Svizzera Italiana (USI), Lugano, Switzerland

Introduction

The advent in recent years of high-performance medical imaging tools, such as ultrasonography (US), computed tomography (CT), magnetic resonance (MRI), and positron emission tomography/computed tomography (PET/CT) with various tracers such as fluorine-18-fluorodeoxy-glucose

(^{18}F-FDG), radiolabelled choline, and radiolabelled prostate-specific membrane antigen, has improved the management of patients [1, 2]. However, using these new imaging modalities has led to the frequent detection of unexpected asymptomatic lesions, "incidentalomas": an increasingly important topic in clinical practice [3].

Because of the high incidence of nodular thyroid disease in the general population, the identification of thyroid incidentalomas (TI) occurs frequently in clinical practice [4, 5]. According to evidence-based data, the prevalence of ^{18}F-FDG TI is about 2–3% of all PET/CTs, approximately two in three showing focal uptake [5]. Focal ^{18}F-FDG TI is defined as any focal uptake corresponding to a given thyroid nodule. The cancer detection rate in focal ^{18}F-FDG TIs is reported to be up to 35–40% [5–7].

The 2015 American Thyroid Association guidelines for adult patients with thyroid nodules and differentiated thyroid cancer recommend performing fine-needle aspiration cytology (FNA) in all ^{18}F-FDG PET/CT TIs with a sonographically confirmed thyroid nodule >1 cm [8]. Although US is the key diagnostic modality in the initial diagnostic assessment of thyroid nodules, the diagnostic accuracy and performance of thyroid imaging reporting and data systems (TIRADSs) in ^{18}F-FDG TI needs to be further validated [9]. Moreover, significant selection bias may influence the calculation of the risk of malignancy (ROM) of ^{18}F-FDG TIs for several reasons. Firstly, the majority of studies of ^{18}F-FDG TI only report outcomes for nodules undergoing thyroid diagnostic work-up, which represent a minority of most institutional series of patients. Oncological patients with ^{18}F-FDG TI have a low likelihood of thyroid surgery if the comorbid non-thyroidal malignancy is more aggressive and/or of worse prognosis [10]. Secondly, most studies use histopathological assessment as the reference standard for malignant lesions and non-neoplastic FNA cytology as the reference standard for benign lesions as benign lesions are often not operated [11]. Thirdly, ^{18}F-FDG TIs with indeterminate FNA, which is the Bethesda system for reporting thyroid cytopathology [12] subcategories III and IV, without histological diagnoses, are generally not included in statistical analyses of ROM even if the expected ROM is not negligible.

In view of the above, before undertaking FNA in an ^{18}F-FDG TI, it is important to know the relative frequency of the BETHESDA cytological subcategories in an ^{18}F-FDG TI. For all thyroid nodules a recently published meta-analysis by Vuong et al. [13]. shows pooled BETHESDA frequencies as follows: Category I non-diagnostic 12.2%, Category II benign 62.3%, Category III AUS/FLUS 8.0%, Category IV follicular neoplasm/suspicious for follicular neoplasm 6.1%, Category V suspicious for malignancy 3.7%, and Category VI malignant 7.4%. The purpose of the current study was to ask the question—are the cytologic findings in ^{18}F-FDG TI different from those in non-FDG PET/CT detected thyroid nodules? It is known for example that Hürthle cell/oncocytic thyroid lesions (HTL) may be over-represented in ^{18}F-FDG avid thyroid nodules [14]. With this information, the clinician can better manage ^{18}F-FDG TI patients using FNA, particularly for HTL. Both benign and malignant HTLs are known to be ^{18}F-FDG avid [15]. However, HTL usually falls into class IV of BETHESDA [12], where the resection rate was reported 60.5% and ROM 28.9% [13].

This study was designed to achieve evidence-based information on the relative distribution of ^{18}F-FDG TIs in the various cytological categories of BETHESDA (non-diagnostic, benign, atypia of undetermined significance or follicular lesion of undetermined significance, follicular neoplasm/suspicious of a follicular neoplasm, suspicious for malignancy, and malignant) [12], providing information useful for the clinical management of ^{18}F-FDG TI.

Methods

Guidelines followed

In this study, all procedures utilized were consistent with PRISMA guidelines [16].

Search strategy

Three investigators (LS, AP, and PT) independently conducted a comprehensive literature search of online databases MEDLINE (PubMed) and Scopus using the following search terms and their combinations: thyroid, nodule, incidentaloma, FDG, PET, positron. A commencement date limit was not used. The last search was carried out on 31 October 2020. No language restrictions were imposed. The search was restricted to human studies. Three investigators (LS, AP, and PT) screened independently titles and abstracts of the retrieved articles, reviewed the full-texts, and then selected articles for inclusion. References from included studies were also screened for additional articles.

Eligibility criteria

The major inclusion criterion was original studies reporting ^{18}F-FDG TIs undergoing FNA and classified according to BETHESDA. The following studies were excluded: (1) with an overlapping patient or nodule data; (2) reporting only some BETHESDA cytologic subcategories (because these results did not allow calculation of the frequency of each cytological subclass of BETHESDA); (3) with ≤10 ^{18}F-FDG TIs. Three researchers (LS, AP, and PT) applied the above criteria

selecting studies for inclusion. Disagreements were resolved via online consensus discussion among all the authors.

Data extraction

For the included studies, the following data were extracted independently and coded in duplicate by three investigators (LS, AP, and PT), in the pilot form: (a) author, publication year, country, study design; (b) number of PET scans performed during the study period; (c) patients' ages and gender; (d) SUVmax value; (e) size of ^{18}F-FDG TIs; (f) number of ^{18}F-FDG TI undergoing FNA; (g) number of ^{18}F-FDG TI in all six BETHESDA categories; (h) number of ^{18}F-FDG TI with a histological diagnosis. The collated details were cross-checked and any discrepancies were fully reconciled by joint re-evaluation among the authors.

Study quality assessment

The risk of bias of the included studies was assessed independently by three investigators (LS, AP, and PT) using the National Heart, Lung, and Blood Institute Quality Assessment Tool (https://www.nhlbi.nih.gov/health-topics/study-quality-assessment-tools).

Data analysis

The characteristics of included studies were summarized. A proportion meta-analysis calculation was used to obtain the pooled rate of ^{18}F-FDG TI assessed in all BETHESDA categories [12]. Six separate meta-analyses were performed to obtain the pooled prevalence (95% confidence interval, 95% CI) of ^{18}F-FDG TIs in the six different BETHESDA categories. Heterogeneity between studies was assessed by using I^2, with 50% or higher values regarded as high heterogeneity. If the presence of heterogeneity was identified, further analyses were performed to explain it. The Egger's test was carried out to evaluate the possible presence of significant publication bias. For statistical pooling of data, a random-effect model was used. A $p < 0.05$ was regarded as significant. All analyses were performed using StatsDirect statistical software (StatsDirect Ltd; Birkenhead, Merseyside, UK).

Results

Study selection

Literature searches using the above algorithms yielded 407 studies. All the records were assessed as depicted in Fig. 1. Of these, 289 were screened, 84 were assessed as

eligible, and 15 [17–31] were included in the final systematic review and meta-analysis.

Study quality assessment

Table 1 summarizes the quality assessment of the 15 included studies. The risk of bias for each study was judged as low for 12 items. All studies were high risk with respect to sample size. None of the studies reported power or sample size justification.

## Qualitative analysis (systematic review)

Table 2 summarizes the main features of the 15 included studies. The 15 studies were published between 2012 and 2020. All studies but one [29] were reported in the English language. Nine studies were performed by Asian/Eastern [17–20, 22–24, 27, 31] and six by Western [21, 25, 26, 28–30] authors. They were single-center or observational cohort studies. The total number of ^{18}F-FDG PET/CT examinations was 232,568 identifying 4031 (1.7%) focal ^{18}F-FDG TIs. The female-to-male ratio was approximately 3:1. Almost all patients were submitted for ^{18}F-FDG PET/CT to stage non-thyroidal cancer. The FNA data according to BETHESDA was available in 2304 ^{18}F-FDG TIs. The final histological diagnoses were reported in 41.3% of cases (618 cancers and 335 benign nodules).

Quantitative analysis (meta-analysis)

The distribution of the 2304 ^{18}F-FDG TIs according to BETHESDA was evaluated. Table 3 shows the pooled prevalence results. The most frequent cytologic category was benign (BETHESDA II) comprising 45% of ^{18}F-FDG TIs. The malignant category (BETHESDA VI) was the second most common cytologic subcategory (19%). Inconsistency was found in all six categories. Publication bias was present in two categories.

In an attempt to explain the above heterogeneity, further analyses were performed. One sensitivity analysis included only those studies with more than 100 TIs [18, 19, 22, 23, 25, 27]. This analysis included 1904 ^{18}F-FDG TIs. The prevalence in the BETHESDA subcategories was substantially unchanged (I 12%, II 47%, III 6%, IV 4%, V 6%, and VI 21%) and the inconsistency remained high (detailed data not shown). A second analysis was performed considering separately the nine studies from Asian/Eastern countries [17–20, 22–24, 27, 31] and the six studies from Western authors [21, 25, 26, 28–30] as shown in Table 4. There was a significantly different prevalence in the BETHESDA category IV between Asian/Eastern and Western studies. With this sub-analysis, no heterogeneity was identified in category IV of Asian/Eastern studies and

Fig. 1 Diagram of the flow of searching data. US ultrasound, CT computed tomography, MRI magnetic resonance imaging, pts patients, FNA fine-needle aspiration. Bethesda system for reporting thyroid cytopathology (BETHESDA)
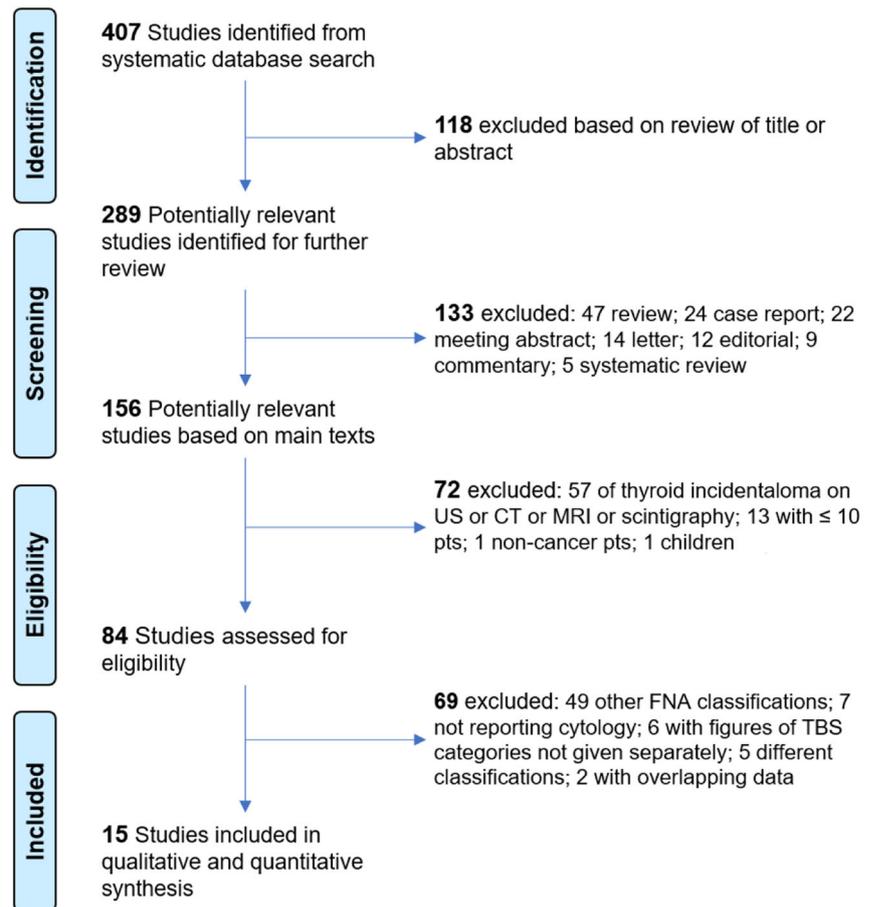
in categories II, III, and V of the Western studies. Unfortunately, data on the frequency and presence or absence of HTL was not available to perform further specific analysis. Forest and funnel plots are included as supplemental material.

Discussion

An evidence-based review detailing information on BETHESDA FNA subcategory outcomes in ^{18}F-FDG focally avid nodules is not currently available in the literature. This systematic review with meta-analysis provides additional evidence-based data on the distribution of ^{18}F-FDG TIs across the full range of BETHESDA cytologic subcategories, enabling more detailed consideration of clinical management in patients undergoing cytologic assessment for ^{18}F-FDG TIs [8].

This study shows that FDG avid nodules show a relative excess of BETHESDA category IV and V FNA (see Table 4). Although the published studies included in this meta-analysis do not give specific information on the prevalence of Hurthle cell neoplasms, the finding of relatively higher frequencies of category IV and V FNA in

thyroid TI implies that this is due to the underlying higher clinical ROM of PET avid thyroid nodules, in combination with a relative excess of HCN in nodules that are FDG avid which would typically fall in BETHESDA categories III and IV [14, 32–34].

Diagnostic assessment of ^{18}F-FDG TIs represents a major challenge in clinical practice. The majority of focal ^{18}F-FDG PET/CT avid thyroid nodules are benign. Unfortunately, we could not derive conclusions about potential relationships between SUVmax values and BETHESDA categories due to the paucity of data. Patients undergoing thyroid ^{18}F-FDG PET/CT generally show more aggressive non-thyroidal comorbid tumors, hence investigation of a potential co-existent thyroid carcinoma may not be a therapeutic priority [10]. Although up to 35–40% of ^{18}F-FDG TIs are malignant, this data is likely to be biased due to the inclusion of patients with non-thyroid tumors [10].^{18}F-FDG-avid thyroid cancer is a malignancy said to be more aggressive than non^{18}F-FDG-avid thyroid carcinoma [35].

This study does not address cancer prevalence among focal ^{18}F-FDG TIs, rather it describes how a cytologic report can influence the clinical management of clinically suspected ^{18}F-FDG TI patients as this information is important, especially for patients with more aggressive

Table 1 Risk of bias summary: review authors' judgements on the quality of the 15 included studies

First author	Lee [17]	Kim [18]	Choi [19]	Lee [20]	Jamsek [21]	Kim [22]	Kim [23]	Yoon [24]	Hagenimana [25]	Li [26]	Suh [27]	Thuillier [28]	De Guevara [29]	Kaliszewski [30]	Kamakshi [31]
1. Research question.	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L
2. Study population.	L	L	L	L	L	L	L	L	L	H	L	L	L	L	L
3. Participation rate.	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
4. Groups recruiting.	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L
5. Sample size.	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H
6. Exposure assessment	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L
7. Sufficient timeframe to see an effect.	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L
8. Different levels of exposure of interest.	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L
9. Exposure measures.	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L
10. Repeated exposure assessment.	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L
11. Outcome measures.	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L
12. Blinding of outcome assessors.	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L
13. Follow-up rate.	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L
14. Statistical analyses.	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L

L low risk, *H* high risk, *NR* not reported

Table 2 Main features of the 15 studies included in the present systematic review

First author [ref]	Year	Country	Total PET/CT performed	TIs	Age	F/M	SUVmax	Size (mm)	FNA classified according to Bethesda							Histology		
									Tot	I	II	III	IV	V	VI	Tot	Mal	Ben
Lee [17]	2012	Korea	327	17	56±10	NA	4.3	NA	16	0	10	1	2	0	3	4	4	0
Kim [18]	2013	Korea	22,674	483	59±11	NA	NA	NA	280	33	155	5	8	7	72	72	71	1
Choi [19]	2014	Korea	7914	171	56±10	84/48	6.3	16	171	33	62	0	4	15	57	68	65	3
Lee [20]	2014	Korea	2368	64	58.8±6.2	24/5	4.5	20	29	2	13	2	0	8	4	15	9	6
Jamsek [21]	2015	Slovenia	2840	148	61.2±12.9	NA	4.2	NA	52	7	24	3	8	3	7	NA	NA	NA
Kim [22]	2015	Korea	18,172	374	58±11	116/40	4	13	235	17	153	13	4	12	36	NA	NA	NA
Kim [23]	2015	Korea	23,462	493	50±10	157/43	5.4	NA	200	14	116	12	6	13	39	72	49	23
Yoon [24]	2015	Korea	NA	87	59±12	48/39	6.2	16.5	87	1	42	4	1	3	36	27	27	0
Hagenimana [25]	2017	Canada	40,914	304	NA	NA	NA	NA	159	36	63	14	23	15	8	46	25	21
Li [26]	2017	USA	NA	20	NA	12/8	NA	NA	20	0	6	2	8	2	2	11	9	2
Suh [27]	2017	Korea	96,942	1360	59±11	1035/325	NA	NA	859	82	229	222	18	31	277	570	325	245
Thuillier [28]	2017	France	10,118	131	64.2±11.6	55/45	NA	NA	60	16	24	5	12	1	2	19	10	9
De Guevara [29]	2020	Chile	5100	119	62	NA	NA	NA	58	1	20	4	4	8	21	NA	NA	NA
Kaliszewski [30]	2020	Poland	NA	56	56±15	42/7	NA	NA	49	0	23	7	12	5	2	49	24	25
Kamakshi [31]	2020	India	1737	204	51	199/89	4.5	23	29	10	11	3	2	1	2	NA	NA	NA
TOT			232,568	4031		1772/649			2304	252	951	296	113	123	569	953	618	335

Age, SUVmax, and size findings were reported in the studies as mean or median

TI thyroid incidentaloma at PET/CT, *F* female, *M* male, *NA* not available

non-thyroid tumors. The results of this study show pooled rates for BETHESDA I non-diagnostic of 10%, BETHESDA II benign 45%, BETHESDA III AUS/FLUS 8%, BETHESDA IV FN/SFN 5%, BETHESDA V SFM 6%, and BETHESDA VI malignant of 19%. The frequency of BETHESDA V suspicious for malignancy and BETHESDA VI malignant FNA in ^{18}F-FDG-avid thyroid nodules is much higher than the 7.4% published rate for Bethesda VI FNA and 3.7% for Bethesda V FNA shown in a recent meta-analysis of Western and Asian patients [13]. This study, therefore, indicates that ^{18}F-FDG TIs have a higher ROM than that observed in the general population of thyroid nodules so requiring further investigation. A 45% of ^{18}F-FDG TIs were cytologically assessed as benign, enabling the use of FNA as a rule-out test in these patients, most of whom have a more aggressive non-thyroidal cancer.

Pooled prevalence data shows that overall, nearly two-thirds of ^{18}F-FDG TIs have either a malignant or benign FNA report, which provides reassurance because the false positive and false negative rates of these categories are negligible (i.e., 2–3% and 0–3%, respectively) [12]. Overall, combining results from Western and Asian/Eastern studies, 13% of ^{18}F-FDG TIs are indeterminate (8% BETHESDA III and 5% BETHESDA IV) although over one-quarter of the patients will have an inconclusive FNA based on Western studies whereas the figure for Asian/Eastern patients is much lower. Because the ROM for the indeterminate FNA categories approaches 30% [13], this represents a challenge for clinical practice. A relatively high number of BETHESDA IV FNA among patients with ^{18}F-FDG TIs are oncocytic/Hürthle cell lesions [36, 37] although most studies do not separately record these lesions in published series.

Ultrasound assessment is known to be suboptimal for detection of follicular and oncocytic thyroid carcinomas as compared to papillary thyroid carcinoma; yet some thyroid cancers with poor prognosis, e.g., some follicular and oncocytic carcinomas, are typically classified as BETHESDA III or IV [38, 39]. Moreover, ^{18}F-FDG avid thyroid cancer is a tumor that is generally more biologically aggressive [40–42]. The possibility of an indeterminate FNA result should be taken into account when requesting FNA in ^{18}F-FDG TI. In addition, the presence of non-invasive follicular thyroid neoplasm with papillary-like nuclear features in Bethesda III and IV categories, as well as in the other categories, should also be considered [43, 44]. High heterogeneity was found in these findings. However, this heterogeneity can be explained according to the

Table 3 Pooled prevalence of the 2304 ^{18}F-FDG TIs in the six Bethesda categories

TBS category	Prevalence (%), 95% CI	I^2 (%)	Egger test (p)
I (ND)	10, 6–14	86.0	0.30
II (B)	45, 37–53	92.7	0.15
III (AUS/FLUS)	8, 3–13	94.6	0.11
IV (FN/SFN)	8, 5–12	88.2	0.001
V (SFM)	6, 4–9	69.8	0.03
VI (M)	19, 13–25	91.6	0.86

TBS the Bethesda system for reporting thyroid cytopathology, ND nondiagnostic, B benign, AUS/FLUS atypia of undetermined significance/follicular lesion of undetermined significance, FN/SFN follicular neoplasms/suspicious for follicular neoplasms, M malignant, SFM suspicious for malignancy, CI confidence interval, I^2 inconsistency

Table 4 Pooled prevalence of the ^{18}F-FDG TIs in the six Bethesda categories in eastern and western studies compared with meta-analysis results for all thyroid nodules by Vuong et al. [13]

	Eastern studies			Western studies		
Studies, n (ref.)	9 [17–20, 22–24, 27, 31]		Vuong et al. [13]	6 [21, 25, 26, 28–30]		Vuong et al. [13]
Nodules, n	1906		69,907	398		75,159
TBS category	Prevalence %, 95% CI	I^2	Prevalence %, 95% CI	Prevalence %, 95% CI	I^2	Prevalence %, 95% CI
I (ND)	10, 6–14	82.2	12.6, 6.7–18.5	9, 2–20	90.4	11.9, 9.1–14.7
II (B)	48, 36–60	95.8	59.8, 51.6–67.9	40, 36–45	0	64.2, 60.0–68.4
III (AUS/FLUS)	7, 1–15	96.9	8.4, 5.5–11.4	9, 6–12	0	7.7, 5.1–10.2
IV (FN/SFN)[a]	2, 2–3	0.3	3.5, 1.9–5.1	19, 12–27	68.9	7.9, 5.7–10.1
V (SFM)	6, 4–8	70.4	4.3, 2.6–6.1	9, 5–13	37.1	3.3, 2.6–4.1
VI (M)	24, 18–31	86.4	10.9, 7.1–14.7	11, 4–21	86.2	4.9, 3.8–6.0

Here, the prevalence is pooled and differs from absolute prevalence since the figures also depend on the weight given to each study

TBS the Bethesda system for reporting thyroid cytopathology, ND nondiagnostic, B benign, AUS/FLUS atypia of undetermined significance/follicular lesion of undetermined significance, FN/SFN follicular neoplasms/suspicious for follicular neoplasms, M malignant, SFM suspicious for malignancy, CI confidence interval

[a]Indicates the categories in which there was a significantly different prevalence between Eastern and Western studies (i.e., the 95% CIs were not overlapping)

sub-analysis performed separating the studies published by Asian/Eastern and Western authors. As shown in Table 4, the heterogeneity was canceled in some cases, as there was a significant difference between Asian/Eastern and Western studies in BETHESDA category IV. The latter finding corroborates the results obtained by Vuong et al. [13]. However, in the Vuong et al. meta-analysis [13] the prevalence of category IV was 7.9% in Western and 3.5% in Asian/Eastern studies, while here we found 19% and 2%, respectively. The reasons for the published differences between Western and Asian/Eastern cytopathology practice are unclear. There may also be differences in nuclear thresholds for papillary thyroid carcinoma although data on this is lacking. Beyond any consideration of the ^{18}F-FDG avid thyroid nodule risk (clinical information, biochemical tests, and US features), we should always keep in mind the context in which we are moving. Indeed, the extrathyroidal PET/CT findings should be carefully considered prior to the decision to undertake thyroid FNA. In this context, there are probably two optimal imaging scenarios to perform FNA biopsy: patients with complete remission of their non-thyroidal cancer and patients with ^{18}F-FDG PET/CT findings suspected of a new diagnosis of metastatic thyroid cancer. In other terms, the higher the risk of finding thyroid cancer and the higher the likelihood of detection of a highly aggressive primary thyroid malignancy, the more appropriate the indication for FNA.

The pooled data of this meta-analysis could be affected by various biases which should be discussed. The studies included in this systematic review recorded 4031 focal ^{18}F-FDG TIs while they reported the results of 2304 (57.1%) FNAs. It is unclear whether patients were managed according to specific clinical features, US-related risk, SUVmax value, or other characteristics associated with the cancers which indicated PET/CT (i.e., selection bias). Whether the cytopathologists were influenced by the FNA indication (i.e., ^{18}F-FDG TI in an oncological patient) is also not reported. The studies report the results of a retrospective review of single-center series of patients. The PET/CT systems utilized were different with different sensitivity and resolutions. Finally, a high statistical heterogeneity among the included studies was found although this heterogeneity can be partially explained after a sub-analysis of the two groups of Asian/Eastern and Western studies.

Conclusions

This meta-analysis shows for the first time that two in three of ^{18}F-FDG TIs undergoing FNA have a malignant (BETHESDA VI) or benign (BETHESDA II) cytology. The remaining approximately one-quarter of cases have an indeterminate FNA and the remainder are non-diagnostic FNA. A significant difference between Asian/Eastern and Western studies is present in the prevalence of the indeterminate category IV. Thyroidologists should be aware of this data to enable better management of patients, especially when an aggressive non-thyroid cancer is present. This evidence-based data suggests guiding clinical decision-making according to the patient's clinical context, including the indication for FNA and extra-thyroidal findings of ^{18}F-FDG PET/CT.

Compliance with ethical standards

Conflict of interest The authors declare no competing interests.

References

1. F. Bertagna, D. Albano, L. Giovanella, R. Giubbini, G. Treglia, F18-choline/C11-choline PET/CT thyroid incidentalomas. Endocrine **64**(2), 203–208 (2019). https://doi.org/10.1007/s12020-019-01841-z

2. F. Bertagna, D. Albano, L. Giovanella, M. Bonacina, R. Durmo, R. Giubbini, G. Treglia, ^{68}Ga-PSMA PET thyroid incidentalomas. Hormones **18**(2), 145–149 (2019). https://doi.org/10.1007/s42000-019-00106-8

3. J.W. O'Sullivan, T. Muntinga, S. Grigg, J.P.A. Ioannidis, Prevalence and outcomes of incidental imaging findings: umbrella review. BMJ **18**(361), k2387 (2018). https://doi.org/10.1136/bmj.k2387

4. J. Jin, C.R. McHenry, Thyroid incidentaloma. Best Pract. Res. Clin. Endocrinol. Metab. **26**(1), 83–96 (2012). https://doi.org/10.1016/j.beem.2011.06.004

5. G. Treglia, F. Bertagna, R. Sadeghi, F.A. Verburg, L. Ceriani, L. Giovanella, L. Focal, Thyroid incidental uptake detected by 18F-fluorodeoxyglucose positron emission tomography. Meta-analysis

on prevalence and malignancy risk. Nuklearmedizin **52**(4), 130–136 (2013). https://doi.org/10.3413/Nukmed-0568-13-03

6. F. Bertagna, G. Treglia, A. Piccardo, R. Giubbini, Diagnostic and clinical significance of F-18-FDG-PET/CT thyroid incidentalomas. J. Clin. Endocrinol. Metab. **97**(11), 3866–3875 (2012). https://doi.org/10.1210/jc.2012-2390

7. G. Signore, D. Albano, L. Giovanella, F. Bertagna, G. Treglia, Evidence-based data about prevalence and risk of malignancy of thyroid incidentalomas detected by different PET radio-pharmaceuticals. Curr. Radiopharm. **13**(2), 89–93 (2020). https://doi.org/10.2174/1874471012666191212115732

8. B.R. Haugen, E.K. Alexander, K.C. Bible, G.M. Doherty, S.J. Mandel, Y.E. Nikiforov, F. Pacini, G.W. Randolph, A.M. Sawka, M. Schlumberger, K.G. Schuff, S.I. Sherman, J.A. Sosa, D.L. Steward, R.M. Tuttle, L. Wartofsky, 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), 1–133 (2016). https://doi.org/10.1089/thy.2015.0020

9. P. Trimboli, L. Knappe, G. Treglia, T. Ruberto, A. Piccardo, L. Ceriani, G. Paone, L. Giovanella, FNA indication according to ACR-TIRADS, EU-TIRADS and K-TIRADS in thyroid incidentalomas at ^{18}F-FDG PET/CT. J. Endocrinol. Investig. **43**(11), 1607–1612 (2020). https://doi.org/10.1007/s40618-020-01244-2

10. D.A. Pattison, M. Bozin, A. Gorelik, M.S. Hofman, R.J. Hicks, A. Skandarajah, ^{18}F-FDG-avid thyroid incidentalomas: the importance of contextual interpretation. J. Nucl. Med. **59**(5), 749–755 (2018). https://doi.org/10.2967/jnumed.117.198085

11. S. Nayan, J. Ramakrishna, M.K. Gupta, The proportion of malignancy in incidental thyroid lesions on 18-FDG PET study: a systematic review and meta-analysis. Otolaryngol. Head Neck Surg. **151**(2), 190–200 (2014). https://doi.org/10.1177/0194599814530861

12. S. Z. Ali, E. Cibas. *The Bethesda System for Reporting Thyroid Cytopathology. Definitions, Criteria, and Explanatory Notes*, 2nd edn. (Springer, 2018).

13. H.G. Vuong, H.T.T. Ngo, A. Bychkov, C.K. Jung, T.H. Vu, K.B. Lu, K. Kakudo, T. Kondo, Differences in surgical resection rate and risk of malignancy in thyroid cytopathology practice between Western and Asian countries: a systematic review and meta-analysis. Cancer Cytopathol. **128**(4), 238–249 (2020). https://doi.org/10.1002/cncy.22228

14. D.N. Poller, H. Megadmi, M.J.A. Ward, P. Trimboli, Hürthle cells on fine-needle aspiration cytology are important for risk assessment of focally PET/CT FDG avid thyroid nodules. Cancers **12**(12), E3544 (2020). https://doi.org/10.3390/cancers12123544

15. D.A. Pryma, M. Schöder, H. Gönen, R.J. Robbins, S.M. Larson, H.W.D. Yeung, Diagnostic accuracy and prognostic value of 18F-FDG PET in Hrthle cell thyroid cancer patients. J Nucl Med **47**(8), 1260–1266 (2006)

16. R. DerSimonian, N. Laird, Meta-analysis in clinical trials. Control Clin. Trials **7**, 177–188 (1986). https://doi.org/10.1016/0197-2456(86)90046-2

17. W.M. Lee, B.J. Kim, M.H. Kim, S.C. Choi, S.Y. Ryu, I. Lim, K. Kim, Characteristics of thyroid incidentalomas detected by pre-treatment [F]FDG PET or PET/CT in patients with cervical cancer. J. Gynecol. Oncol. **23**(1), 43–47 (2012). https://doi.org/10.3802/jgo.2012.23.1.43

18. H. Kim, S.J. Kim, I.J. Kim, K. Kim, Thyroid incidentalomas on FDG PET/CT in patients with non-thyroid cancer - a large retrospective monocentric study. Onkologie **36**(5), 260–264 (2013). https://doi.org/10.1159/000350305

19. J.S. Choi, Y. Choi, E.K. Kim, J.H. Yoon, J.H. Youk, K.H. Han, H.J. Moon, W.J. Kang, J.Y. Kwak, A risk-adapted approach using US features and FNA results in the management of thyroid

20. incidentalomas identified by 18F-FDG PET. Ultraschall Med. **35**(1), 51–58 (2014). https://doi.org/10.1055/s-0033-1335328

20. S. Lee, T. Park, S. Park, K. Pahk, S. Rhee, J. Cho, E. Jeong, S. Kim, J.G. Choe, The clinical role of dual-time-point (18)F-FDG PET/CT in differential diagnosis of the thyroid incidentaloma. Nucl. Med. Mol. Imaging **48**(2), 121–129 (2014). https://doi.org/10.1007/s13139-013-0247-z

21. J. Jamsek, I. Zagar, S. Gaberscek, M. Grmek, Thyroid lesions incidentally detected by (18)F-FDG PET-CT—a two centre retrospective study. Radiol. Oncol. **49**(2), 121–127 (2015). https://doi.org/10.2478/raon-2014-0039

22. B.H. Kim, S.J. Kim, K. Kim, H. Kim, S.J. Kim, W.J. Kim, Y.K. Jeon, S.S. Kim, Y.K. Kim, I.J. Kim, High metabolic tumor volume and total lesion glycolysis are associated with lateral lymph node metastasis in patients with incidentally detected thyroid carcinoma. Ann. Nucl. Med. **29**(8), 721–729 (2015). https://doi.org/10.1007/s12149-015-0994-2

23. S.J. Kim, S. Chang, Predictive value of intratumoral heterogeneity of F-18 FDG uptake for characterization of thyroid nodules according to Bethesda categories of fine needle aspiration biopsy results. Endocrine **50**(3), 681–688 (2015). https://doi.org/10.1007/s12020-015-0620-z

24. J.H. Yoon, A. Cho, H.S. Lee, E.K. Kim, H.J. Moon, J.Y. Kwak, Thyroid incidentalomas detected on 18F-fluorodeoxyglucose-positron emission tomography/computed tomography: thyroid imaging reporting and data system (TIRADS) in the diagnosis and management of patients. Surgery **158**(5), 1314–1322 (2015). https://doi.org/10.1016/j.surg.2015.03.017

25. N. Hagenimana, J. Dallaire, E. Vallée, M. Belzile, Thyroid incidentalomas on 18FDG-PET/CT: a metabolico-pathological correlation. J. Otolaryngol. Head Neck Surg. **46**(1), 22 (2017). https://doi.org/10.1186/s40463-017-0200-8

26. Y. Li, M. Cui, N. Azar, D. Nakamoto, C.W. Michael, Cytological evaluation by fine needle aspiration biopsy of incidental focal increased fluorodeoxyglucose uptake in thyroid on positron emission tomography scan. Diagn. Cytopathol. **45**(6), 501–506 (2017). https://doi.org/10.1002/dc.23695

27. C.H. Suh, Y.J. Choi, J.J. Lee, W.H. Shim, J.H. Baek, H.C. Chung, Y.K. Shong, D.E. Song, T.Y. Sung, J.H. Lee, Comparison of core-needle biopsy and fine-needle aspiration for evaluating thyroid incidentalomas detected by ^{18}F-fluorodeoxyglucose positron emission tomography/computed tomography: a propensity score analysis. Thyroid **27**(10), 1258–1266 (2017). https://doi.org/10.1089/thy.2017.0192

28. P. Thuillier, N. Roudaut, G. Crouzeix, M. Cavarec, P. Robin, R. Abgral, V. Kerlan, P.Y. Salaun, Malignancy rate of focal thyroid incidentaloma detected by FDG PET-CT: results of a prospective cohort study. Endocr. Connect. **6**(6), 413–421 (2017). https://doi.org/10.1530/EC-17-0099

29. H.D. Ladrón de Guevara, M.C. Munizaga, S.N. García, B.C. Letelier, F.A.Wash, Frecuencia de malignidad en incidentalomas tiroideos detectados con tomografía por emisión de positrones/tomografía computada (PET/CT) con F18-FDG de cuerpo entero [frequency of malignancy in thyroid incidentalomas detected by whole body 18F-FDG PET/CT]. Rev Med Chil **148**(1), 10–16 (2020). https://doi.org/10.4067/S0034-98872020000100010

30. K. Kaliszewski, D. Diakowska, M. Ziętek, B. Knychalski, M. Aporowicz, K. Sutkowski, B. Wojtczak, Thyroid incidentaloma as a "PAIN" phenomenon—does it always require surgery? Medicine **97**(49), e13339 (2020). https://doi.org/10.1097/MD.0000000000013339

31. K. Kamakshi, A. Krishnamurthy, V. Karthik, P. Vinodkumar, R. K. Kumar, K.M. Lakshmipathy, Positron emission tomography-computed tomography-associated incidental neoplasms of the thyroid gland. World J. Nucl. Med. **19**(1), 36–40 (2020). https://doi.org/10.4103/wjnm.WJNM_33_19

32. Y. Ren, N. Kyriazidis, W.C. Faquin, S. Soylu, D. Kamani, R. Saade, N. Torchia, C. Lubitz, L. Davies, N. Stathatos, A.E. Stephen, G.W. Randolph, The presence of Hürthle cells does not increase the risk of malignancy in most Bethesda categories in thyroid fine-needle aspirates. Thyroid **30**(3), 425–431 (2020). https://doi.org/10.1089/thy.2019.0190

33. K. S. Wong, T. E. Angell, J. A. Barletta, J. F. Krane. Hürthle cell lesions of the thyroid: progress made and challenges remaining. Cancer Cytopathol. Oct (2020). https://doi.org/10.1002/cncy.22375.

34. K.S. Wong, V.Y. Jo, A.C. Lowe, W.C. Faquin, A.A. Renshaw, A. A. Shah, M.H. Roh, E.B. Stelow, J.F. Krane, Malignancy risk for solitary and multiple nodules in Hürthle cell-predominant thyroid fine-needle aspirations: a multi-institutional study. Cancer Cytopathol. **128**(1), 68–75 (2020). https://doi.org/10.1002/cncy.22213

35. A. Piccardo, P. Trimboli, L. Foppiani, G. Treglia, G. Ferrarazzo, M. Massollo, G. Bottoni, L.Giovanella, PET/CT in thyroid nodule and differentiated thyroid cancer patients. The evidence-based state of the art. Rev. Endocr. Metab. Disord. **20**(1), 47–64 (2019). https://doi.org/10.1007/s11154-019-09491-2

36. K.A. Pathak, T. Klonisch, R.W. Nason, W.D. Leslie, FDG-PET characteristics of Hürthle cell and follicular adenomas. Ann. Nucl. Med. **30**(7), 506–509 (2016). https://doi.org/10.1007/s12149-016-1087-6

37. S. Zandieh, W. Pokieser, P. Knoll, C. Sonneck-Koenne, M. Kudlacek, S. Mirzaei, Oncocytic adenomas of thyroid-mimicking benign or metastatic disease on 18F-FDG-PET scan. Acta Radiol. **56**(6), 709–713 (2015). https://doi.org/10.1177/0284185114537928

38. M. Castellana, A. Piccardo, C. Virili, L. Scappaticcio, G. Grani, C. Durante, L. Giovanella, P. Trimboli, Can ultrasound systems for risk stratification of thyroid nodules identify follicular carcinoma? Cancer Cytopathol. **128**(4), 250–259 (2020). https://doi.org/10.1002/cncy.22235

39. P. Trimboli, M. Castellana, A. Piccardo, F. Romanelli, G. Grani, L. Giovanella, C. Durante. The ultrasound risk stratification systems for thyroid nodule have been evaluated against papillary carcinoma. A meta-analysis. Rev. Endocr. Metab. Disord. Sep. https://doi.org/10.1007/s11154-020-09592-3 (2020).

40. C. Are, J.F. Hsu, R.A. Ghossein, H. Schoder, J.P. Shah, A.R. Shaha, Histological aggressiveness of fluorodeoxyglucose positronemission tomogram (FDG-PET)-detected incidental thyroid carcinomas. Ann. Surg. Oncol. **14**, 3210–3215 (2007). https://doi.org/10.1245/s10434-007-9531-4

41. A. Piccardo, M. Puntoni, F. Bertagna, G. Treglia, L. Foppiani, F. Arecco, R. Giubbini, M. Naseri, A. Cistaro, M. Cabria, F. Bardesono, L. Ceriani, F. Orlandi, L. Giovanella, ^{18}F-FDG uptake as a prognostic variable in primary differentiated thyroid cancer incidentally detected by PET/CT: a multicentre study. Eur. J. Nucl. Med. Mol. Imaging **41**(8), 1482–1491 (2014). https://doi.org/10.1007/s00259-014-2774-y

42. G. Treglia, L. Giovanella, Prognostic role of FDG-PET/CT in differentiated thyroid carcinoma: where are we now? J. Med. Imaging Radiat. Oncol. **59**(3), 278–280 (2015). https://doi.org/10.1111/1754-9485.12317

43. M. Bongiovanni, L. Giovanella, F. Romanelli, P. Trimboli, Cytological diagnoses associated with noninvasive follicular thyroid neoplasms with papillary-like nuclear features according to the Bethesda system for reporting thyroid cytopathology: a systematic review and meta-analysis. Thyroid **29**(2), 222–228 (2019). https://doi.org/10.1089/thy.2018.0394

44. M. Bongiovanni, W.C. Faquin, L. Giovanella, C. Durante, P. Kopp, P. Trimboli, Impact of non-invasive follicular thyroid neoplasms with papillary-like nuclear features (NIFTP) on risk of malignancy in patients undergoing lobectomy/thyroidectomy for suspected malignancy or malignant fine-needle aspiration cytology findings: a systematic review and meta-analysis. Eur. J. Endocrinol. **181**(4), 389–396 (2019). https://doi.org/10.1530/EJE-19-0223