Review Article on Asian and Western Practice in Thyroid Pathology: Similarities and Differences

Atypia of undetermined significance/follicular lesion of undetermined significance: Asian *vs.* non-Asian practice, and the Singapore experience

Li Yin Ooi, Min En Nga

Department of Pathology, National University Hospital, Singapore

Contributions: (I) Conception and design: All authors; (II) Administrative support: None; (III) Provision of study materials or patients: None; (IV) Collection and assembly of data: All authors; (V) Data analysis and interpretation: All authors; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Min En Nga, MBBS, MRCPath, FRCPath, FIAC. Department of Pathology, National University Hospital, 5, Lower Kent Ridge Road, 1 Main Building, Level 3, Singapore 119074. Email: patngame@nus.edu.sg.

Abstract: The Bethesda System for Reporting Thyroid Cytopathology has paved the way for comparisons of the practice of thyroid cytology in many different regions. However, there have been comparatively few studies documenting differences between Asian and non-Asian practice. Here, we aim to compare a few key parameters between the two regions, focusing on the indeterminate category of atypia of undetermined significance (AUS)/follicular lesion of undetermined significance (FLUS). We compared its incidence, resection rates (RRs), risk of malignancy (ROM), rate of repeat fine needle aspiration (rFNA), ROMs of cytomorphologic subcategories of nuclear atypia (AUS-N) *vs.* architectural atypia (AUS-A), and, finally, the incidence of papillary thyroid carcinoma (PTC) *vs.* follicular neoplasms (FNs) in resected AUS/FLUS cases in Asian and non-Asian regions. Where possible, these metrics were compared with the Singapore experience from a tertiary referral institution. While the incidence of AUS/FLUS was similar in both regions, we found geographical differences in the RRs and ROMs, which may reflect a higher collective threshold for surgery in Asian countries. However, both cohorts showed higher ROMs in the AUS-N subcategory as compared to the AUS-A subcategory, supporting the subclassification of the AUS/FLUS based on the presence of nuclear atypia. We also observed a higher incidence of AUS-N coupled with a higher incidence of PTC in resected AUS/FLUS nodules in Asian cohorts, while AUS-A and follicular-patterned neoplasms featured more prominently in the non-Asian cohorts. These incidences may account for the starkly different molecular approaches that we noted—in Asian (chiefly Korean and Chinese) centers, *BRAF* mutational analysis was favored, while gene panels and gene expression classifiers were more frequently applied in non-Asian centers (chiefly in the United States of America). Overall, the data from Singapore appears more closely aligned to non-Asian trends, despite its geographical location in Southeast Asia and its predominantly Asian population. We conclude that there is significant heterogeneity in the outcomes of the AUS/FLUS categories between and within regions, which is only partially explained by regional variations, and may also reflect different regional diagnostic and management practices. This highlights the importance of understanding the local context in the interpretation of indeterminate Bethesda categories, rather than adopting a "one-size fits all" approach.

Keywords: Bethesda classification; fine needle aspiration cytology; atypia of undetermined significance (AUS); follicular lesion of undetermined significance (FLUS); thyroid

Submitted Mar 30, 2020. Accepted for publication Jun 30, 2020.
doi: 10.21037/gs-20-555
View this article at: http://dx.doi.org/10.21037/gs-20-555

Introduction

It is an undisputed fact that regional differences exist in diagnostic and management practices between different geographic regions. Much of this may be related to epidemiological variations, however, a combination of cultural factors and national and regional healthcare infrastructure does play a role.

In the recent literature, there has been some documentation of the differences in diagnostic thresholds and/or management of thyroid nodules between the East and West (1,2). For example, Bychkov *et al.* has demonstrated that the impact of non-invasive follicular thyroid neoplasm with papillary-like nuclear features (NIFTP) on Bethesda category outcomes is less marked in Asia than in the West (2,3). The appreciation of regional differences is highly relevant in current global practice because of the recent worldwide movement towards applying international consensus guidelines for cytologic terminology in many different organ systems, with the Bethesda System for Reporting Thyroid Cytopathology (TBSRTC) being one of the pioneer consensus systems in non-gynecologic cytology (4).

One of the most challenging Bethesda categories is the atypical category—atypia of undetermined significance (AUS)/follicular lesion of undetermined significance (FLUS). Not only is there diversity in cytomorphologic subgroups within this category, some diagnostic criteria may overlap with those of other Bethesda categories. Indeed, a number of independent studies have sought to address the differential risk of malignancy (ROM) associated with the presence of nuclear atypia within this category in both Asian and non-Asian study groups (5-16). Understandably, the AUS/FLUS diagnosis is also heavily influenced by individual and institutional experience, and bears an inherent element of subjectivity in interpretation. This is reflected in the wide variation in both the reported incidence rates and malignancy rates in this category. Recently, Vuong *et al.* demonstrated regional differences in the resection rates (RRs) and ROM of AUS/FLUS, suggesting that Asian clinicians as a whole tend to adopt a more conservative approach in their management of AUS/FLUS nodules as compared to their Western counterparts (1).

In view of its diagnostic and therapeutic heterogeneity, we draw focus to the AUS/FLUS category, comparing the practice of thyroid cytopathology between Asian and non-Asian countries, with a particular emphasis on the Singapore experience. We aim to characterize the differences between Asian and non-Asian practice in terms of its incidence, RRs, rates of repeat fine needle aspiration (rFNA), and ROMs. In addition, we review the outcomes of repeat FNA and the differential ROMs amongst cytomorphologic subgroups—cases with nuclear atypia *vs.* architectural atypia only. Where possible, we will also compare the prevalence of well-differentiated neoplasms amongst surgically resected cases.

Lastly, we will comment on the different practices in molecular interrogation of AUS/FLUS nodules between Asian and non-Asian countries. For this comparison, we draw our information from studies that focus on AUS/FLUS outcomes rather than primarily on molecular testing.

Methods and materials

Identification and selection of studies

We searched for relevant articles in PubMed from January 2007 to August 2019 using the search term "(fine needle OR cytology) AND thyroid AND Bethesda". Study titles and abstracts were screened for candidate articles. Studies were included if they fulfilled the following criteria: (I) reporting of thyroid FNA results using TBSRTC 2009 or 2017, and (II) adequate data for AUS/FLUS cases provided for either of the following parameters of practice:

(I) Incidence;
(II) RR, ROM and overall risk of malignancy (OROM);
(III) Rate of repeat FNA (rFNA) and rate of a more definitive Bethesda category on rFNA;
(IV) RR without rFNA, ROM without rFNA, ROM with rFNA;
(V) ROM for the AUS/FLUS subgroups of architectural atypia (AUS-A) and nuclear atypia (AUS-N). Any cases with nuclear atypia (regardless of whether architectural atypia was present) were included in the AUS-N cohort. Subgroups with Hürthle cell change and atypical lymphoid cells were not examined.

The exclusion criteria for the ROM analyses were: (I) studies specifying the use of other risk predictors such as molecular or ultrasound findings to select for surgery, (II) studies using non-consecutive cases, (III) studies that used core needle biopsy as an intermediate form of assessment, (IV) studies that used non-surgical follow-up to determine outcomes, (V) case reports, (VI) reviews and meta-analyses.

Full-text screening and data extraction

The full-text of candidate articles was screened and data extracted into a predefined dataset. The following data points from all the included studies were extracted: year of publication, institution, city, country, period of study. For each parameter below, the following data points were extracted:

(I) Incidence: number of FNAs performed and number of AUS/FLUS cases;

(II) RR, ROM and OROM: number of AUS/FLUS cases, number of surgical resections and number of malignancies diagnosed on histology. Malignancy inferred from non-surgical outcomes (i.e., core needle biopsy or imaging) were not included. NIFTP and tumors of uncertain malignant potential were included in the calculation of malignancy. Papillary microcarcinomas were only included in the malignant follow-up category if they were specifically correlated with the nodule that underwent FNA. Studies with less than 30 surgical resections were excluded from calculations of RR, ROM and OROM;

(III) Rate of rFNA and rate of a more definitive Bethesda category on repeat FNA: Number of AUS/FLUS, number of rFNA, and number of rFNA yielding a more definitive Bethesda category. "More definitive Bethesda category" was defined as Bethesda categories with a more definite management pathway, i.e., benign, follicular neoplasm (FN), suspicious for malignancy (SM) or malignant;

(IV) RR without rFNA, ROM without rFNA, ROM with rFNA: number of AUS/FLUS, number of surgical resections without repeat FNA, number of malignancies on resection histology without repeat FNA, number of surgical resections after repeat FNA, and number of malignancies on resection histology after repeat FNA;

(V) ROM of AUS/FLUS subgroups: number of AUS-A and AUS-N, number of surgical resections of AUS-A and AUS-N, and number of malignancies on resection in AUS-A and AUS-N. Where reported, the incidence of the two cytomorphological subgroups and the final histologic diagnoses were noted. Studies with less than 30 surgical resections were excluded from calculations of ROM.

For studies that contained potentially overlapping cases from the same institution, the study with the larger number of FNAs was selected. For studies that compared pre- and post-Bethesda prevalence, only cases that were classified post-Bethesda were included. For studies with cases from both Asian and non-Asian institutions, the cases were extracted and analyzed separately according to the geographical region.

Additionally, studies were also screened for molecular testing and specific histologic outcomes. The following data were reviewed where available: (I) type of molecular tests performed and their results, (II) in studies with at least 50 surgically resected cases, the proportion of papillary thyroid carcinoma (PTC), follicular variant PTC (FVPTC), and FNs, including follicular adenoma (FA) and follicular carcinoma (FC), among surgically resected cases.

Data analysis

The following definitions were applied: incidence of AUS/FLUS is equivalent to the proportion of AUS/FLUS cases among total number of FNAs performed. The RR is equivalent to the proportion of cases in a specified group that underwent surgical resection. The ROM is equivalent to the proportion of malignant cases confirmed by histopathological examination among the surgically resected cases of a specified group. The OROM is equivalent to the proportion of malignant cases confirmed by histopathological examination among all the cases in a specified group (both with and without resection). The rate of rFNA is equivalent to the proportion of AUS/FLUS cases that received one or more repeat FNAs.

Statistical analysis was performed using the statistical software JAMOVI (https://www.jamovi.org), which is built in the R statistical language. Each parameter was independently evaluated with meta-analysis of proportion using DeSimonian-Laird method and 95% CIs were pooled using a random-effect model. The comparison of proportions between Asian and Non-Asian studies were performed using subgroup analysis.

Assessment of publication or small study bias

To assess the presence of publication bias and small study bias, funnel plots of effect estimates from individual studies were performed. Funnel plot asymmetry was determined using the rank correlation test and regression

test. A P value of <0.05 was considered statistically significant for the presence of funnel plot asymmetry, which indicates the presence of publication or small study bias.

Results

We identified 859 papers for title and abstract screening and included 210 for full text review. One hundred and twenty studies were eventually selected for extraction of data. Asian series included those from South-East Asia (Singapore and Thailand), East Asia (mainland China, Taiwan, Japan and Korea), South Asia (India and Bangladesh), and West Asia (Kingdom of Saudi Arabia, United Arab Emirates, Turkey, Israel, and Egypt). Non-Asian series included those from the United States of America (USA), Europe (France, Poland, Spain, Finland, Macedonia and Czech Republic), Canada, Brazil, South Africa and Australia.

Incidence of AUS/FLUS

Forty-six Asian series and 43 non-Asian series were included in the pooled analysis of AUS/FLUS incidence (*Table 1*). The total number of cases amongst the Asian and non-Asian cohorts were 260,169 and 149,460 in this analysis. The incidence of the diagnosis of AUS/FLUS was not found to be significantly different between the Asian and non-Asian cohorts [8.8% (95% CI, 7.4–10.2%) *vs.* 9.1% (95% CI, 7.9–10.3%), P=0.69]. The incidence of AUS/FLUS in a single tertiary referral center in Singapore was 6.4% (16).

RR, ROM and OROM within AUS/FLUS

Twenty-eight Asian series and 35 non-Asian series were included in the pooled analysis of RR, ROM and OROM (*Table 1*). This yielded a total of 19,232 cases in the Asian studies and 10,567 in the non-Asian studies. There were statistically significant differences in both the RRs and ROMs. The RR in the Asian cohort was significantly lower than that of the non-Asian cohort [33.2% (95% CI, 27.7–38.8%) *vs.* 43.4% (95% CI, 36.7–50.2%), P=0.02] (*Figure 1A,B*). Conversely, the ROM in the Asian cohort was higher than that of the non-Asian cohort [43.2% (95% CI, 32.4–54.0%) *vs.* 26.8% (95% CI, 23.3–30.3%), P=0.005] (*Figure 2A,B*).

The OROM was not significantly different between the Asian and non-Asian cohorts [13.0% (95% CI, 10.6–15.4%) *vs.* 11.4% (95% CI, 9.2–13.7%), P=0.32].

In Singapore, the RR in a single institution was 44.3%, the ROM was 27.0% and the OROM was 12.0%.

Repeat FNA in AUS/FLUS and outcomes

Nine Asian and 16 non-Asian series were included in the pooled analysis of the rate of rFNAs after an initial FNA diagnosis of AUS/FLUS, and the outcomes of the rFNA (*Table 2*). In total, there were 7,541 cases amongst the Asian studies and 4,479 cases in the non-Asian studies. There were no regional differences in the rate of rFNA [Asian, 29.4% (95% CI, 22.7–36.2%); non-Asian, 28.0% (95% CI, 22.5–33.5%), P=0.74] or the rate of a more definitive Bethesda category on rFNA [Asian, 65.9% (95% CI, 49.3–82.5%); non-Asian, 61.2% (95% CI, 55.6–66.7%), P=0.62].

In the Singapore series, the rate of rFNA was 23.6%, whilst the rate of a more definitive Bethesda category on rFNA was 67.1%.

RR and ROM for direct to surgery cases; ROM after repeat FNA

Eight Asian and 11 non-Asian series were included in the pooled analysis of the RR and ROM without repeat FNA (i.e., cases that went directly to surgery) and ROM after rFNA (*Table 2*). There were in total 8,083 cases in the Asian cohort and 3,914 cases in the non-Asian cohort.

In the direct to surgery cohorts, although the Asian group showed a trend towards a lower RR compared to the non-Asian cohort, this was not statistically significant [23.9% (95% CI, 18.0–29.8%) *vs.* 30.7% (95% CI, 21.1–40.2%), P=0.23]. In the Singapore series, the RR in directly resected cases was 40.5%.

The ROMs in both direct to surgery and post-rFNA cohorts paralleled each other, trending towards higher ROMs in the Asian groups, but the differences did not achieve statistical significance. The ROM in directly resected cases was 34.1% (95% CI, 18.2–49.9%) in the Asian cohort and 23.7% (95% CI, 16.1–31.4%) in the non-Asian cohort; P=0.25. The ROM of cases that were resected after rFNA was 40.2% (95% CI, 14.0–66.4%) in the Asian cohort and 28.1% (95% CI, 19.6–36.6%) in the non-Asian cohort; P=0.49.

In the Singapore series, the ROM in the direct to surgery cohort was 26.4%, and the ROM after rFNA was 33.3%.

Table 1 Studies reviewed for AUS/FLUS incidence, RR, ROM and OROM

Study	Region	Study period	No. of FNAs	AUS/FLUS						
				No. of total cases	I	R	M	RR	ROM	OROM
Asian										
Gan 2017, (16)	Singapore	2008–2014	4,815	309	0.064	137	37	0.443	0.270	0.120
Keelawat 2017, (17)	Thailand	2010–2017	7,447	159	0.021	117	42	0.736	0.359	0.264
Thewjitcharoen 2019, (18)	Thailand	2010–2017	2,735	128	0.047	30	3	0.234	0.100	0.023
Mao 2017, (19)	Mainland China	2014–2015	3,090	442	0.143	121	43	0.274	0.355	0.097
Zheng 2018, (20)	Mainland China	2014–2016	7,355	736	0.100	NA	NA	NA	NA	NA
Ke 2019, (21)	Mainland China	2011–2016	13,351	1,764	0.132	157	83	0.089	0.529	0.047
Jan 2019, (22)	Taiwan	2012–2016	29,937	770	0.026	367	204	0.477	0.556	0.265
Satoh 2017, (23)	Japan	2015–2016	1,600	171	0.107	47	7	0.275	0.149	0.041
Kim 2011, (24)	Korea	2007–2009	865	141	0.163	33	24	0.234	0.727	0.170
Chung 2011, (25)	Korea	2005–2010	3,962	515	0.130	166	108	0.322	0.651	0.210
Hyeon 2014, (26)	Korea	2011–2012	6,402	551	0.086	231	157	0.419	0.680	0.285
Park 2014, (7)	Korea	2010–2011	3,589	331	0.092	95	77	0.287	0.811	0.233
Yoo 2015, (27)	Korea	2010–2012	11,988	772	0.064	287	234	0.372	0.815	0.303
Jung 2015, (28)	Korea	2011–2014	18,091	163	0.009	71	47	0.436	0.662	0.288
Koh 2016, (29)	Korea	2012–2013	1,754	123	0.070	32	28	0.260	0.875	0.228
Kim SJ 2017, (30)	Korea	2012	5,321	346	0.065	NA	NA	NA	NA	NA
Kim SD 2017, (31)	Korea	2010–2014	8,458	660	0.078	NA	NA	NA	NA	NA
Kim M 2017, (32)	Korea	2011	42,132	4,100	0.097	828	569	0.202	0.687	0.139
Hong 2018, (33)	Korea	2011–2014	6,365	717	0.113	96	70	0.134	0.729	0.098
Mondal 2013, (34)	India	2009–2012	1,020	10	0.010	NA	NA	NA	NA	NA
Mehra 2015, (35)	India	2010–2012	225	11	0.049	NA	NA	NA	NA	NA
Garg 2015, (36)	India	2012–2014	100	4	0.040	NA	NA	NA	NA	NA
Arul 2015, (37)	India	2012–2015	603	60	0.100	41	10	0.683	0.244	0.167
Mahajan 2017, (38)	India	2010–2015	4,562	116	0.025	NA	NA	NA	NA	NA
Kumari 2019, (39)	India	NA	1,050	10	0.010	NA	NA	NA	NA	NA
Naz 2014, (40)	Bangladesh	NA	528	67	0.127	NA	NA	NA	NA	NA
Ozluk 2011, (41)	Turkey	2004–2007	581	25	0.043	NA	NA	NA	NA	NA
Ustün 2012, (42)	Turkey	2007–2011	14,629	3,903	0.267	1,756	228	0.450	0.130	0.058
Firat 2012, (43)	Turkey	2010–2012	764	75	0.098	NA	NA	NA	NA	NA
Dincer 2013, (15)	Turkey	2009–2010	7,658	368	0.048	88	23	0.239	0.261	0.063

Table 1 (*continued*)

Table 1 (*continued*)

Study	Region	Study period	No. of FNAs	AUS/FLUS						
				No. of total cases	I	R	M	RR	ROM	OROM
Tepeoğlu 2014, (44)	Turkey	2009–2011	1,021	100	0.098	79	10	0.790	0.127	0.100
Onder 2014, (13)	Turkey	2009–2012	6,310	421	0.067	103	18	0.245	0.175	0.043
Muratli 2014, (45)	Turkey	2008–2013	1,607	140	0.087	NA	NA	NA	NA	NA
Gocun 2014, (46)	Turkey	2010–2013	4,916	347	0.071	82	22	0.236	0.268	0.063
Kuru 2016, (47)	Turkey	2011–2015	5,157	607	0.118	179	41	0.295	0.229	0.068
Selek 2016, (48)	Turkey	2009–2014	10,769	560	0.052	112	46	0.200	0.411	0.082
Turkyilmaz 2017, (49)	Turkey	2011–2015	9,938	1,007	0.101	305	111	0.303	0.364	0.110
Öcal 2019, (50)	Turkey	2009–2015	233	106	0.455	NA	NA	NA	NA	NA
Mufti 2012, (51)	KSA	2005–2010	250	2	0.008	NA	NA	NA	NA	NA
Al-Abbadi 2013, (52)	KSA	2010–2011	205	15	0.073	NA	NA	NA	NA	NA
Alabdulqader 2015, (53)	KSA	2012–2013	251	25	0.100	NA	NA	NA	NA	NA
Al Dawish 2017, (54)	KSA	2012–2014	1,433	131	0.091	42	6	0.321	0.143	0.046
Hirsch 2015, (55)	Israel	2011–2012	3,927	457	0.116	66	30	0.144	0.455	0.066
Ronen 2019, (56)	Israel	2013–2017	287	38	0.132	NA	NA	NA	NA	NA
Sinna 2012, (57)	Egypt	2005–2010	296	40	0.135	NA	NA	NA	NA	NA
Al-Abbadi 2017, (58)	UAE	2010–2014	2,592	115	0.044	33	11	0.287	0.333	0.096
Non-Asian										
Jo 2010, (59)	USA	1992–2009	3,080	104	0.034	53	9	0.510	0.170	0.087
Rabaglia 2010, (60)	USA	2008–2009	765	91	0.119	32	4	0.352	0.125	0.044
Luu 2011, (8)	USA	2004–2009	7,072	222	0.031	127	33	0.572	0.260	0.149
Olson 2011, (61)	USA	2009–2011	3,956	388	0.098	133	43	0.343	0.323	0.111
VanderLaan 2011, (6)	USA	2005–2009	4,691	512	0.109	199	96	0.389	0.482	0.188
Bongiovanni 2012, (62)	USA	2007–2009	3,724	248	0.067	132	19	0.532	0.144	0.077
Chen 2012, (63)	USA	2006–2011	393	61	0.155	32	6	0.525	0.188	0.098
Horne 2012, (64)	USA	2008–2009	6,205	171	0.028	106	29	0.620	0.274	0.170
Harvey 2013, (65)	USA	2009–2011	3,432	72	0.021	31	6	0.431	0.194	0.083
Nagarkatti 2013, (66)	USA	2005–2007	5,391	254	0.047	151	24	0.594	0.159	0.094
Olson 2013, (67)	USA	2009–2012	3,885	575	0.148	106	32	0.184	0.302	0.056
Theoharis 2013, (68)	USA	2007–2008	3,207	95	0.030	NA	NA	NA	NA	NA
Broome 2014, (69)	USA	2009–2012	3,200	306	0.096	170	28	0.556	0.165	0.092
Ho 2014, (70)	USA	2008–2011	8,862	709	0.080	381	144	0.537	0.378	0.203
Lee 2014, (71)	USA	2011–2012	NA	122	NA	60	16	0.492	0.267	0.131
Mathur 2014, (72)	USA	2009–2013	4,827	806	0.167	255	99	0.316	0.388	0.123

Table 1 (*continued*)

Table 1 (*continued*)

Study	Region	Study period	No. of FNAs	AUS/FLUS						
				No. of total cases	I	R	M	RR	ROM	OROM
McElroy 2014, (73)	USA	2006	97	7	0.072	NA	NA	NA	NA	NA
Sullivan 2014, (74)	USA	2003–2012	5,665	332	0.059	168	56	0.506	0.333	0.169
Walts 2014, (75)	USA	2008–2012	5,079	457	0.090	181	60	0.396	0.331	0.131
Wu 2014, (12)	USA	2002–2008	3,346	670	0.200	138	25	0.206	0.181	0.037
Deniwar 2015, (76)	USA	2010–2014	723	94	0.130	65	22	0.691	0.338	0.234
Kantola 2016, (77)	USA	2011–2013	2,156	159	0.074	67	12	0.421	0.179	0.075
Brandler 2016, (78)	USA	2012–2014	11,481	976	0.085	321	112	0.329	0.349	0.115
Krauss 2016, (79)	USA	1999–2013	5,574	238	0.043	55	14	0.231	0.255	0.059
Shrestha 2016, (80)	USA	2006–2012	NA	221	NA	101	29	0.457	0.287	0.131
Valderrabano 2016, (81)	USA	2008–2014	2,829	340	0.120	188	53	0.553	0.282	0.156
Guo 2017, (82)	USA	2011–2015	236	8	0.034	NA	NA	NA	NA	NA
Deaver 2018, (83)	USA	2011–2015	2,019	231	0.114	94	23	0.407	0.245	0.100
Seagrove-Guffey 2018, (84)	USA	2015	893	43	0.048	NA	NA	NA	NA	NA
Bresler 2019, (85)	USA	2012–2015	2,258	213	0.094	99	37	0.465	0.374	0.174
Williams 2013, (86)	Canada	2006–2010	1,491	281	0.188	NA	NA	NA	NA	NA
Bernstein 2016, (87)	Canada	2010–2013	1,944	233	0.120	187	86	0.803	0.460	0.369
Erivwo 2018, (88)	Canada	2010–2013	3,285	181	0.055	NA	NA	NA	NA	NA
Rosario 2014, (9)	Brazil	2009–2013	1,742	150	0.086	135	34	0.900	0.252	0.227
Rosario 2017, (89)	Brazil	2015	708	98	0.138	NA	NA	NA	NA	NA
Reuters 2018, (90)	Brazil	2012–2013	980	70	0.071	NA	NA	NA	NA	NA
Mosca 2018, (11)	Brazil	2009–2013	1,093	384	0.351	NA	NA	NA	NA	NA
Firat 2012, (43)	France	2009–2011	2,277	210	0.092	35	6	0.167	0.171	0.029
Ratour 2013, (91)	France	2010–2011	2,210	337	0.152	39	9	0.116	0.231	0.027
Stanek-Widera 2016, (92)	Poland	2010–2016	16,656	395	0.024	35	8	0.089	0.229	0.020
Estrada Muñoz 2017, (93)	Spain	2010–2014	3,032	151	0.050	55	8	0.364	0.145	0.053
Paajanen 2018, (94)	Finland	2011–2012	415	32	0.077	NA	NA	NA	NA	NA
Mileva 2018, (95)	Macedonia	2012–2016	4,738	281	0.059	90	34	0.320	0.378	0.121
Sarkis 2014, (96)	Australia	2010–2013	2,076	97	0.047	54	5	0.557	0.093	0.052
Fatman 2015, (97)	S. Africa	2008–2011	1,767	141	0.080	44	17	0.312	0.386	0.121

AUS/FLUS, atypia of undetermined significance/follicular lesion of undetermined significance; FNA, fine needle aspiration; I, incidence; R, number of resected cases; M, number of malignant cases; RR, resection rate; ROM, risk of malignancy; OROM, overall risk of malignancy; NA, not available.

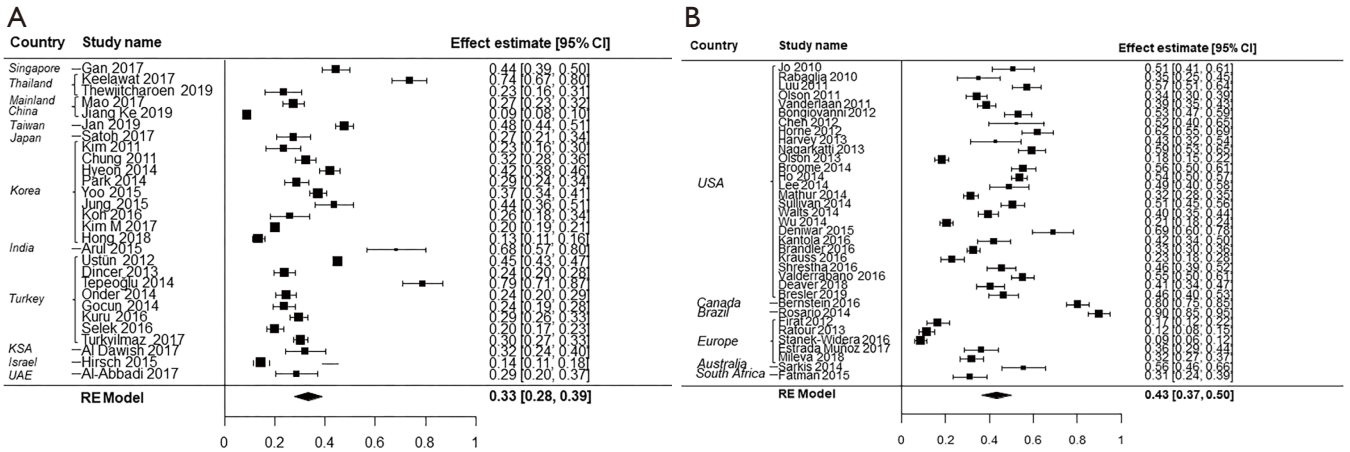

Figure 1 Forest plots of meta-analysis on the resection rates of AUS/FLUS nodules. (A) Asian series; (B) non-Asian series. AUS, atypia of undetermined significance; FLUS, follicular lesion of undetermined significance.

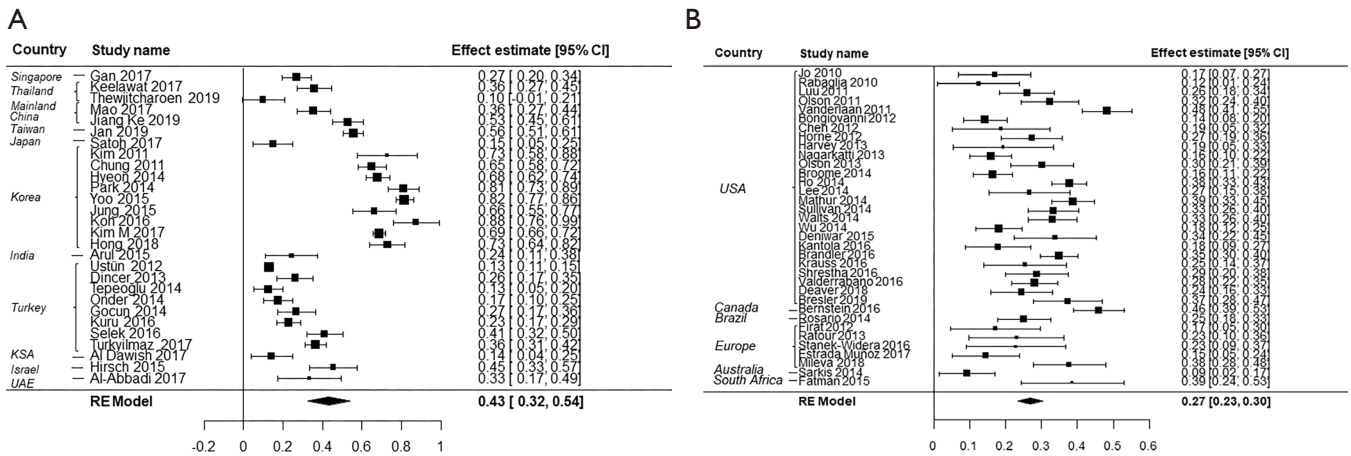

Figure 2 Forest plots of meta-analysis on the risk of malignancy of AUS/FLUS nodules. (A) Asian series; (B) non-Asian series. AUS, atypia of undetermined significance; FLUS, follicular lesion of undetermined significance.

AUS/FLUS with nuclear vs. architectural atypia: ROM and incidence

Eight Asian and 15 non-Asian series were included in the pooled analysis (*Table 3*); these included 1,112 cases in the Asian cohort and 1,563 cases in the non-Asian cohort. The ROM was higher in AUS-N than AUS-A; this association was apparent in both the Asian and the non-Asian series. In the Asian series, the ROMs of AUS-N and AUS-A were 49.6% (95% CI, 25.5–73.7%) and 17.0% (95% CI, 11.1–22.8%) respectively; P=0.049 (*Figure 3A*). This difference is also borne out in the non-Asian series, where the ROM of AUS-N *vs.* AUS-A was 45.8% (95% CI, 39.3–52.4%) *vs.* 24.0% (95% CI, 15.7–32.3%); P<0.001 (*Figure 3B*).

In Singapore, the ROM of AUS-N cases was significantly higher than that in AUS-A cases (36.8% *vs.* 14.7%, P=0.006).

Within the AUS-N cohort, there was no significant regional difference in ROM between Asian and non-Asian groups (P=0.73); and the same finding was obtained in the AUS-A cohort (P=0.43). Amongst the series that substratified AUS/FLUS by the presence of nuclear *vs.* architectural atypia, the actual incidences of AUS-N and AUS-A amongst all AUS/FLUS cases were documented in only four Asian (two from Korea and two from Turkey; n=1,566) and six non-Asian series (three from USA, one from Canada, one from Italy and one from South Africa;

Table 2 Studies reviewed for repeat FNA in AUS/FLUS and outcomes, RR and ROM in direct to surgery cases, and ROM after repeat FNA

Study	Region	No. of Cat III	Direct to Surgery				Repeat FNA						
			No. of cases	M	RR	ROM	No. of cases	No. of MD	R	M	Rate	Rate of MD	ROM
Asian													
Gan 2017, (16)	Singapore	309	125	33	0.41	0.26	73	49	12	4	0.24	0.67	0.33
Jan 2019, (22)	Taiwan	770	254	132	0.33	0.52	NA	NA	113	72	NA	NA	0.64
Hyeon 2014, (26)	Korea	551	NA	NA	NA	NA	274	214	NA	NA	0.50	0.78	NA
Yoo 2015, (27)	Korea	772	142	109	0.18	0.77	243	176	145	125	0.32	0.72	0.86
Ustün 2012, (42)	Turkey	3903	1171	176	0.30	0.15	1,366	458	585	67	0.35	0.34	0.12
Dincer 2013, (15)	Turkey	368	72	21	0.20	0.29	74	52	16	2	0.20	0.70	0.13
Gocun 2014, (46)	Turkey	347	57	13	0.16	0.23	NA	NA	25	9	NA	NA	0.36
Onder 2014, (13)	Turkey	421	NA	NA	NA	NA	116	79	NA	NA	0.28	0.68	NA
Kuru 2016, (47)	Turkey	607	122	19	0.20	0.16	171	NA	57	22	0.28	NA	0.39
Turkyilmaz 2017, (49)	Turkey	1,007	139	48	0.14	0.35	NA	NA	166	63	NA	NA	0.38
Hirsch 2015, (55)	Israel	457	NA	NA	NA	NA	137	89	NA	NA	0.30	0.65	NA
Ronen 2019, (56)	Israel	38	NA	NA	NA	NA	17	7	NA	NA	0.45	0.41	NA
Al-Abbadi 2017, (58)	UAE	115	NA	NA	NA	NA	9	9	NA	NA	0.08	NA	NA
Non-Asian													
Olson 2011, (61)	USA	388	NA	NA	NA	NA	81	43	NA	NA	0.201	0.53	NA
Chen 2012, (63)	USA	61	NA	NA	NA	NA	26	11	NA	NA	0.43	0.42	NA
Nagarkatti 2013, (66)	USA	254	125	20	0.49	0.16	51	28	26	4	0.20	0.55	0.15
Broome 2014, (69)	USA	306	117	18	0.38	0.15	101	52	48	10	0.33	0.52	0.21
Ho 2014, (70)	USA	709	350	135	0.49	0.39	96	54	31	9	0.14	0.56	0.29
Lee 2014, (71)	USA	122	NA	NA	NA	NA	17	11	NA	NA	0.14	0.65	NA
Sullivan 2014, (74)	USA	332	118	40	0.36	0.34	86	48	42	14	0.26	0.56	0.33
Walts 2014, (75)	USA	457	78	20	0.17	0.26	285	NA	103	40	0.62	NA	0.39
Brandler 2016, (78)	USA	976	264	88	0.27	0.33	281	190	57	24	0.29	0.68	0.42
Shrestha 2016, (80)	USA	221	57	15	0.26	0.26	111	68	44	14	0.50	0.61	0.32
Deaver 2018, (83)	USA	231	NA	NA	NA	NA	33	22	NA	NA	0.14	0.67	NA
Seagrove-Guffey 2018, (84)	USA	43	17	5	0.40	0.29	32	10	4	2	0.74	0.31	0.50
Erivwo 2018, (88)	Canada	181	NA	NA	NA	NA	57	28	NA	NA	0.32	0.49	NA
Reuters 2018, (90)	Brazil	70	18	0	0.26	0	11	10	7	3	0.16	0.91	0.43
Ratour 2013, (91)	France	337	NA	NA	NA	NA	86	60	NA	NA	0.26	0.70	NA
Stanek-Widera 2016, (92)	Poland	395	27	6	0.07	0.22	180	NA	8	2	0.46	NA	0.25
Estrada Muñoz 2017, (93)	Spain	151	37	7	0.25	0.19	61	42	18	1	0.40	0.69	0.06
Sarkis 2014, (96)	Australia	97	NA	NA	NA	NA	12	10	NA	NA	0.12	0.83	NA

AUS/FLUS, atypia of undetermined significance/follicular lesion of undetermined significance; rFNA, repeat fine needle aspiration; M, number of malignant cases; RR, resection rate; ROM, risk of malignancy; R, number of resected cases; MD, cases with a more definitive category on repeat FNA; NA, not available.

Table 3 Studies reviewed for incidence and ROM of nuclear *vs.* architectural atypia in AUS/FLUS

Study	Country	Study period	No. of AUS/FLUS	AUS-A					AUS-N				
				No. of cases	R	M	I	ROM	No. of cases	R	M	I	ROM
Asian													
Gan 2017, (16)	Singapore	2008–2014	309	NA	75	11	NA	0.147	NA	57	21	NA	0.368
Chung 2011, (25)	Korea	2005–2010	515	NA	4	2	NA	0.500	NA	92	55	NA	0.598
Ryu 2014, (98)	Korea	2008–2012	NR	NA	61	15	NA	0.246	NA	55	26	NA	0.473
Hyeon 2014, (26)	Korea	2011–2012	551	120	38	5	0.218	0.132	431	193	152	0.782	0.788
Park 2014, (7)	Korea	2010–2011	331	42	9	2	0.127	0.222	214	68	66	0.647	0.971
Dincer 2013, (15)	Turkey	2009–2010	368	85	26	7	0.231	0.269	283	62	16	0.769	0.258
Çuhaci 2014, (14)	Turkey	2010–2012	NR	NA	91	18	NA	0.198	NA	185	45	NA	0.243
Onder 2014, (13)	Turkey	2009–2012	421	134	34	2	0.318	0.059	257	62	16	0.610	0.258
Non-Asian													
Luu 2011, (8)	USA	2004–2009	222	154	79	11	0.694	0.139	68	48	22	0.306	0.458
Olson 2011, (61)	USA	2009–2011	388	NA	30	8	NA	0.267	NA	62	30	NA	0.484
VanderLaan 2011, (6)	USA	2005–2009	512	NA	45	11	NA	0.244	NA	84	40	NA	0.476
Horne 2012, (64)	USA	2008–2009	171	107	25	4	0.626	0.160	64	33	25	0.374	0.758
Ho 2014, (70)	USA	2008–2011	709	NA	78	44	NA	0.564	NA	28	15	NA	0.536
Lee 2014, (71)	USA	2011–2012	122	NA	41	10	NA	0.244	NA	16	4	NA	0.250
Mathur 2014, (72)	USA	2009–2013	806	NA	49	19	NA	0.388	NA	105	57	NA	0.543
Walts 2014, (75)	USA	2008–2012	457	NA	52	7	NA	0.135	NA	75	36	NA	0.480
Wu 2014, (12)	USA	2002–2008	670	NA	32	8	NA	0.250	NA	41	13	NA	0.317
Shrestha 2016, (80)	USA	2006–2012	221	NA	21	2	NA	0.095	NA	66	23	NA	0.348
Valderrabano 2017, (5)	USA	2008–2015	241	105	62	2	0.436	0.030	84	42	17	0.349	0.400
Bernstein 2016, (87)	Canada	2010–2013	233	122	84	34	0.524	0.405	111	103	52	0.476	0.505
Srbova 2015, (99)	Czech	2001–2012	NR	NA	33	3	NA	0.091	NA	43	10	NA	0.233
Rossi 2017, (100)	Italy	2008–2009	269	140	102	27	0.520	0.265	46	40	26	0.171	0.650
Fatman 2015, (97)	S. Africa	2008–2011	141	93	20	8	0.660	0.400	48	24	9	0.340	0.375

AUS/FLUS, atypia of undetermined significance/follicular lesion of undetermined significance; AUS-A, AUS/FLUS with architectural atypia; AUS-N, AUS/FLUS with nuclear atypia; R, number of resected cases; M, number of malignant cases; I, incidence; ROM, risk of malignancy; NA, not available.

n=1,142). Nevertheless, there appeared to be a significant geographical difference in the incidence of AUS-A and AUS-N. The Asian cohort showed a higher incidence of AUS-N than the non-Asian cohort [70.3% (95% CI, 61.8–78.8%) *vs.* 33.5% (95% CI, 24.2–42.7%); P<0.001] (*Figure 4A*). This trend was reversed in AUS-A, where the Asian cohort demonstrated lower rates than the non-Asian cohort [22.3% (95% CI, 14.7–29.8%) *vs.* 57.6% (95% CI, 49.5–65.7%); P<0.001] (*Figure 4B*).

Incidence of papillary carcinoma and follicular neoplasms in surgically resected cases of AUS/FLUS

To further investigate for possible reasons for the regional

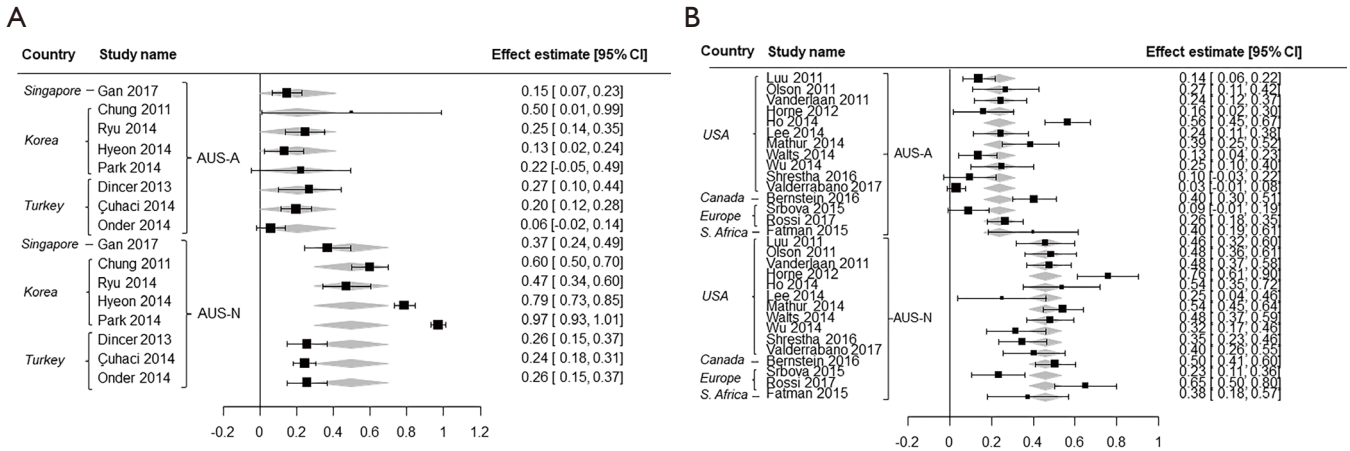

Figure 3 Forest plots of meta-analysis on risk of malignancy of AUS-A *vs.* AUS-N. (A) Asian countries; (B) non-Asian countries. AUS-A, AUS/FLUS with architectural atypia; AUS-N, AUS/FLUS with nuclear atypia; AUS/FLUS, atypia of undetermined significance/follicular lesion of undetermined significance.

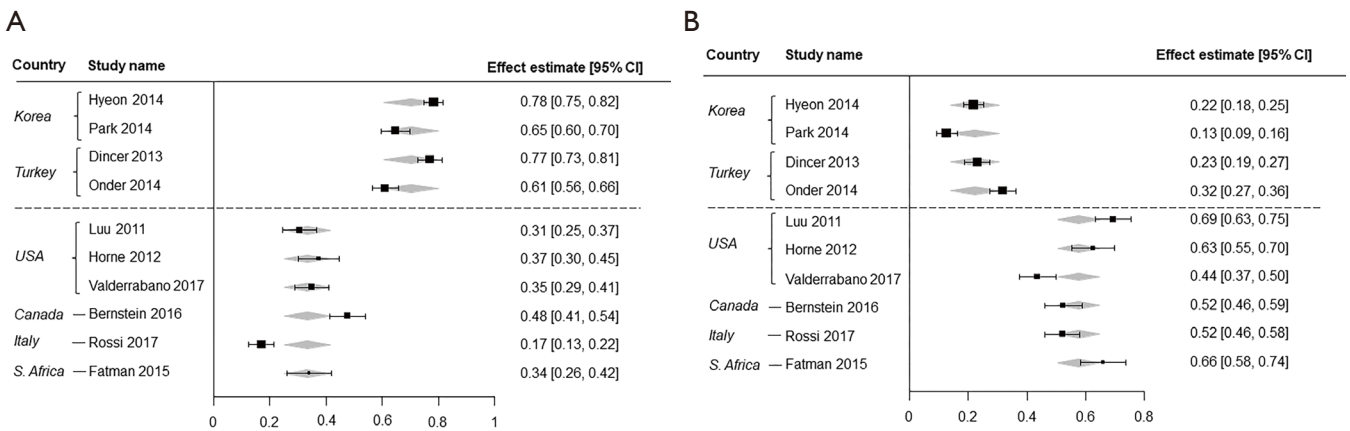

Figure 4 Forest plots of meta-analysis on incidences of AUS subgroups between Asian and non-Asian series. (A) AUS-N incidence; (B) AUS-A incidence. AUS-A, AUS/FLUS with architectural atypia; AUS-N, AUS/FLUS with nuclear atypia; AUS/FLUS, atypia of undetermined significance/follicular lesion of undetermined significance.

differences in the incidences of AUS-A and AUS-N, we reviewed the prevalence of specific tumor subtypes in studies with at least 50 surgically resected cases with documented histologic diagnoses (*Table 4*). Nine Asian studies (one from Singapore, two from mainland China, one from Taiwan, three from Korea, and two from Turkey) and six non-Asian studies were reviewed (three from USA, one from Canada, one from Macedonia and one from Australia); which yielded a cumulative total of 2,318 cases and 1,076 cases in the Asian and non-Asian cohorts respectively. Amongst the Asian studies, the incidence of PTC (of any subtype) ranged from 16.1% to 65.7%, with the Singapore series exhibiting the lowest incidence. Excluding Singapore, the incidence ranged from 31.5% to 65.7%. Amongst the non-Asian studies, the incidence ranged from 9.3% to 45.5%. The pooled incidence of PTC trended higher in the Asian studies than the non-Asian studies [46.3% (95% CI, 34.8–57.9%) *vs.* 29.1% (95% CI, 17.3–40.9%), P=0.061] (*Figure 5A*). On exclusion of the results from the Singapore series, the PTC incidence in Asian *vs.* non-Asian studies was 50.2% *vs.* 29.1% respectively, and this difference reached statistical significance (P=0.009).

Table 4 Studies reviewed for incidence of PTC, FVPTC and FN in surgically resected cases of AUS/FLUS

Study	Country	Study period	No. of resected AUS/FLUS	No. of cases			Incidence		
				PTC	FVPTC	FN	PTC	FVPTC	FN
Asian									
Gan 2017, (16)	Singapore	2008–2014	137	22	6	NA	0.161	0.273	NA
Mao 2017, (19)	Mainland China	2014–2015	121	42	1	8	0.347	0.024	0.066
Ke 2019, (21)	Mainland China	2011–2016	157	79	NA	NA	0.503	NA	NA
Jan 2019, (22)	Taiwan	2012–2016	367	183	NA	43	0.499	NA	0.117
Chung 2011, (25)	Korea	2005–2010	166	109	NA	9	0.657	NA	0.054
Hyeon 2014, (26)	Korea	2011–2012	231	151	9	17	0.654	0.060	0.074
Kim M 2017, (32)	Korea	2011	722	471	NA	77	0.652	NA	0.107
Selek 2016, (48)	Turkey	2009–2014	112	42	NA	NA	0.375	NA	NA
Turkyilmaz 2017, (49)	Turkey	2011–2015	305	96	44	8	0.315	0.458	0.026
Non-Asian									
VanderLaan 2011, (6)	USA	2005–2009	199	86	NA	77	0.432	NA	0.387
Broome 2014, (69)	USA	2009–2012	170	23	11	52	0.135	0.478	0.306
Ho 2014, (70)	USA	2008–2011	381	125	61	69	0.328	0.488	0.181
Bernstein 2016, (87)	Canada	2010–2013	187	85	63	NA	0.455	0.741	NA
Sarkis 2014, (96)	Australia	2010–2013	54	5	NA	19	0.093	NA	0.352
Mileva 2018, (95)	Macedonia	2012–2016	85	26	8	28	0.306	0.308	0.329

AUS/FLUS, atypia of undetermined significance/follicular lesion of undetermined significance; PTC, papillary thyroid carcinoma; FVPTC, follicular variant PTC; FN, follicular neoplasm (including follicular adenoma and follicular carcinoma).

The subtypes of PTC were described in four Asian studies (one from Singapore, one from mainland China, one from Korea and one from Turkey) and four non-Asian studies (two from USA, one from Canada and one from Macedonia). The incidence of FVPTC as a proportion of all the PTCs ranged from 2.4% to 45.8% in Asian studies, whilst the range was 30.8% to 74.1% in non-Asian studies. The pooled incidence of FVPTC as a proportion of all the PTCs was lower in the Asian studies than the non-Asian studies [19.2% (95% CI, 3.8–34.7%) *vs.* 51.4% (95% CI, 32.9–69.8%), P=0.006] (*Figure 5B*).

For FNs (FA, FC), these were specifically mentioned in the histologic outcome in six Asian studies (one from mainland China, one from Taiwan, three from Korea, and one from Turkey); and five non-Asian studies (three from the USA, one from Macedonia and one from Australia) (6,19,22,25,26,32,49,69,70,95,96). Within Asian studies, the range of incidence was 2.6% to 11.7%, compared to 18.1% to 38.7% in the non-Asian studies (*Figure 5C*). The pooled incidence of FNs for Asian *vs.* non-Asian studies was 7.4% (95% CI, 4.0–10.7%) *vs.* 30.7% (95% CI, 21.2–40.1%); P<0.001.

Molecular studies in Asian vs. non-Asian centers

The type of molecular testing performed was noted when documented in the studies reviewed. It should be noted that the studies reviewed did not focus primarily on molecular risk stratification of indeterminate nodules, but, rather, on documenting the follow-up of indeterminate nodules.

Amongst the Asian centers, molecular testing was performed on indeterminate thyroid nodules in eight centers including those in mainland China (20) and in several university centers in Korea (7,24,26-30).

In both mainland China and Korea, the most frequent molecular test performed was mutational testing for *BRAF* (most frequently targeting the *BRAF* V600E mutation). Testing methods included Sanger sequencing

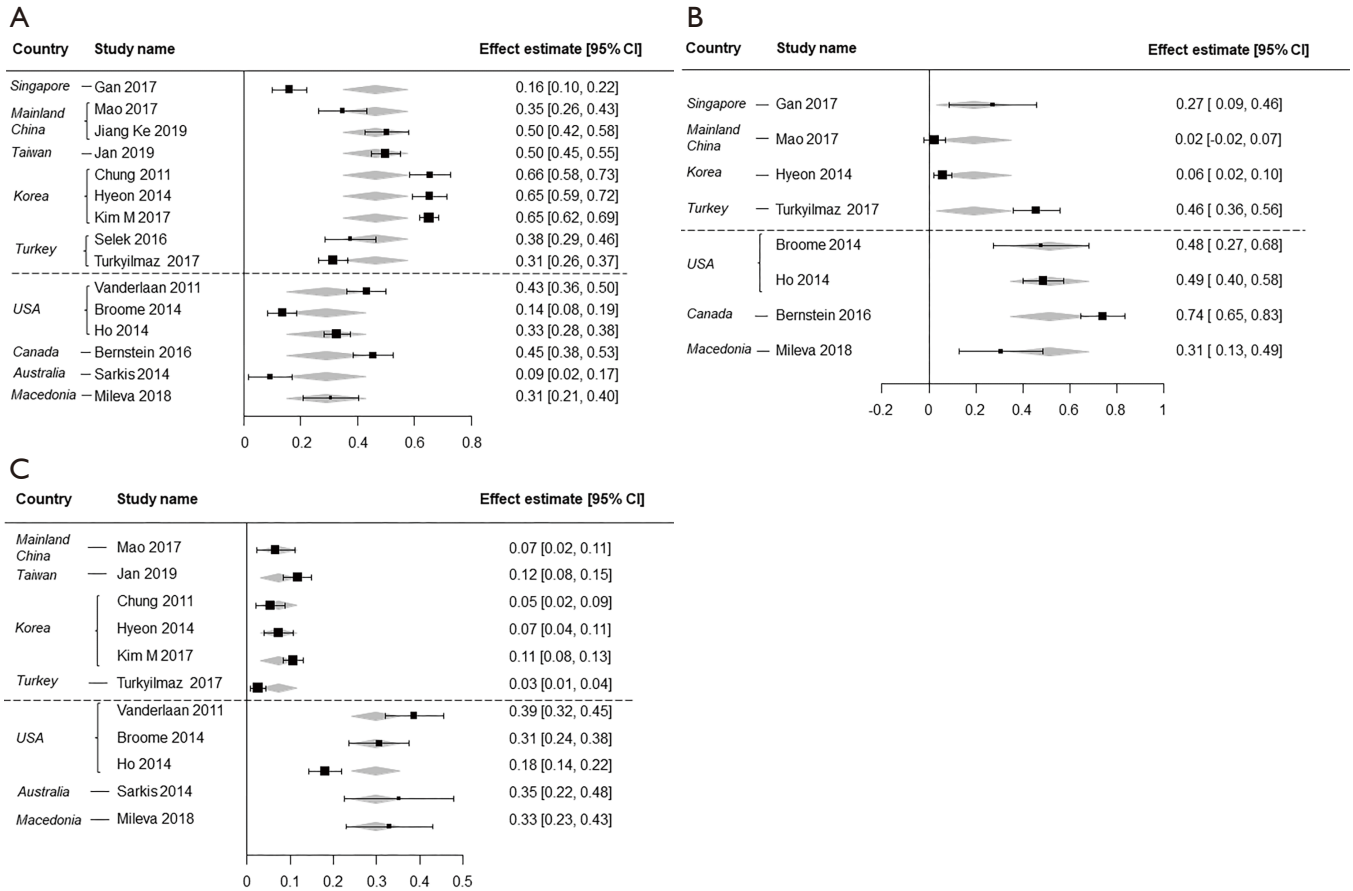

Figure 5 Histologic outcomes in Asian *vs.* non-Asian series. (A) Incidence of PTC; (B) proportion of FVPTC (among PTCs); (C) incidence of FNs. PTC, papillary thyroid carcinoma; FVPTC, follicular variant papillary thyroid carcinoma; FNs, follicular neoplasms.

(both mainland China and Korea), real-time PCR, dual-priming oligonucleotide based multiplex PCR (DPO-PCR), pyrosequencing and PCR-restriction fragment length polymorphism-based analysis (PCR-RFLP). With some of the more sensitive methods of testing, occasional false-positive cases were documented, with Park *et al.* reporting three false positive cases that, on histology, yielded a FA, Hurthle cell adenoma and a Hurthle cell carcinoma (7). Hence, the lowest positive predictive value for the *BRAF* V600E mutation was 90.6%, with the lowest reported specificity at 83.3% (7,20,24,26,28).

Gene array testing or multigene mutation testing was not reported in the Asian studies reviewed.

In the non-Asian papers, molecular tests were described in five studies, all from the USA (74,81-84). In contrast to Asian cohorts, all employed multigene or array-based analysis rather than focused *BRAF* mutation analysis. Tests included the Afirma gene expression classifier (GEC) (74,82,83), miRInform panel (81) and an oncogenic panel which included several *RAS* mutations, *BRAF*, *PIK3CA* and also *RET/PTC* and *PAX8/PPARγ* rearrangements performed by a centralized laboratory (84). None of the studies performed isolated *BRAF* mutation testing. The overall helpfulness of molecular testing in accurate risk stratification for surgery was modest. Seagrove-Guffey *et al.* found the oncogenic panel test unhelpful, with one *NRAS*-mutated case yielding a benign histologic outcome (adenomatoid nodule) whilst a negative case yielded PTC on histology (84). Valderrabano *et al.* tested mirINform on 19 AUS/FLUS cases, and found that none of the three excised cases yielded a positive result, whilst the negative predictive value (NPV) was 84% (81). They reported the results to be "worse than

expected", with a possible reason given that malignancies that were cytologically classified as AUS/FLUS tended to be FVPTC or other follicular-patterned cancers such as FC, which may have resulted in a higher proportion of negative tests. The differential prevalence of follicular-patterned neoplasms between Asian and non-Asian cohorts has been documented in the previous section.

Publication and small study bias

We noted funnel plot asymmetry in the studies included for the pooled analysis of incidence, overall RR and OROM of AUS/FLUS, as well as the incidence of AUS-A and FN (*Figure S1*). As most of the studies reviewed are observational and non-comparative, there were no "undesirable" results or characteristics like significance levels that may have caused publication bias. Hence, we surmise that the funnel plot asymmetry seen in these group of studies is possibly related to small study bias which could lead to exaggeration of effect. In the subset of studies that compared ROMs in cytomorphological subgroups (AUS-A *vs.* AUS-N), funnel plots showed no strong evidence of publication bias (*Figure S2*), with the rank correlation and regression tests confirming the absence of funnel plot asymmetry (P=0.056 and P=0.721).

Discussion

This review describes the geographic differences in the incidences, RRs, ROMs and diagnostic practices in the AUS/FLUS Bethesda category, focusing on Singapore, other Asian countries and non-Asian regions. Whilst meta-analytical methods for evaluating proportions were applied in this work, it should be noted that this is a broad approach that encompasses many aspects of practice. To capture a wider pool of studies, we have included reports that meet the adequacy criteria for the specific parameter under analysis only, recognizing and accepting the possibility of selective reporting.

Incidence of AUS/FLUS and RRs

AUS/FLUS is a challenging diagnostic category bearing a degree of inherent subjectivity. In our analysis, we found that the pooled incidence of AUS/FLUS was comparable between Asian and non-Asian series, at 8.8% and 9.1% respectively, which is close to the recommended prevalence by TBSRTC of less than 10%. In the reference study from Singapore which was conducted in our institution, the data showed an incidence of 6.4%, which is below the pooled Asian incidence. This may reflect the contribution of rapid on-site evaluation (ROSE) by the cytotechnologists as well as on-site provisional reporting in our pathologist-led FNA clinic—services which are widely utilized in our institution, but may not be uniformly practiced worldwide. ROSE serves to optimize specimen collection and processing, thereby reducing preparation artefacts and addressing adequacy issues upfront, both of which can contribute to the incidence of AUS/FLUS. The relatively low incidence may also reflect the reporting of thyroid FNAs by a smaller group of dedicated cytopathologists.

In Singapore, thyroid FNAs are performed mostly in the specialist setting, by clinicians (endocrinologists, surgeons, radiologists) and pathologists. In the reference study from our institution, the practice setting is that of a tertiary referral center with a multiracial patient population (chiefly comprising Chinese, Malays and Indians) with a minor proportion of overseas nationals hailing mostly from the Association of Southeast Asian Nations (ASEAN) countries. Thyroid FNAs are performed under direct palpation or ultrasound guidance. FNA diagnoses are made according to the framework of TBSRTC, with subsequent management guided by a combination of the Bethesda recommendations and clinical/imaging features.

With regard to the management of AUS/FLUS nodules, some regional differences surfaced in our pooled analysis. We found that the overall RR (with and without rFNA) was higher in the non-Asian cohort as compared to Asian series (43.4% *vs.* 33.2%, P=0.02), which is consistent with Vuong *et al.*'s recent meta-analysis (1). Although the overall RR was higher in the non-Asian cohort, we did not find any significant regional differences in the rate of rFNA and the RR of cases that went directly to surgery (i.e., did not undergo rFNA). Although the reasons underpinning the regional differences in RR is beyond the scope of this study, we surmise that the selection threshold for surgery is influenced by multiple considerations that may be region specific, such as cultural beliefs, health behaviors, healthcare costs and accessibility, all of which become particularly important in the management of indeterminate nodules. In addition, rates of surgery may also be influenced by regional variations in malpractice and litigation climates (101,102). The trend towards a more conservative approach in Asian countries may also in part reflect the move towards active surveillance of indeterminate nodules and low-risk carcinomas (103), for example, as documented in the

guidelines from the Japanese Thyroid Association (104).

Interestingly, we found that in our institution in Singapore, the overall RR was 44.3%, which more closely approximated the non-Asian figure of 43.4%. A possible reason for this may be the fact that many of our clinicians undergo higher professional fellowship training in Western regions including the United Kingdom or the USA, and hence may be strongly influenced by Western practices and guidelines. The trend towards a lower threshold for surgery is also reflected in the incidence of the cases that underwent repeat FNA and those that underwent surgery directly. In the Singapore study, the rate of repeat FNA in AUS/FLUS nodules was 23.6%, lower than the pooled incidences of both the Asian (29.4%) and the non-Asian studies (28.0%). Conversely, the proportion of cases that underwent surgery without a repeat FNA was higher in Singapore (40.5%) than the pooled incidences of both the Asian (23.9%) and the non-Asian cohorts (30.7%). A possible reason for the tendency towards surgery could be the specialist setting and the relative accessibility to tertiary surgical services in our institution.

ROM

We found that the ROM was higher in the Asian cohort, at 43.2%, compared to 26.8% in the non-Asian cohort (P=0.005), whilst the OROMs were not statistically different. This geographical difference replicates Vuong *et al.*'s findings (1), and could be related to differences in true disease prevalence, divergent inclusion criteria for AUS/FLUS, discrepant thresholds for surgery or variable thresholds for histopathologic diagnosis of thyroid tumors. Given that we concurrently demonstrated a difference in the collective RR between Asian and non-Asian series, it is likely that the variation in the propensity for surgery accounts significantly for the regional difference in ROM.

With our finding that the RR of AUS nodules in Singapore more closely mirrors the non-Asian practice, it is unsurprising that the ROM of AUS nodules in Singapore is also closer to that of the non-Asian cohort (27% *vs.* 26.8%), and lower than the pooled Asian rate of 43.2%. This is also observed in the ROM in the direct to surgery cohort, which was 26.4%, and closer to the non-Asian rate of 23.7% than the Asian rate of 34.1%. This is likely to be related to the higher proportion of cases that underwent surgery directly without a repeat FNA.

Taking a broader view, in terms of the overall ROM among all Bethesda categories (I to VI), our single-year institutional audit revealed a ROM of 19.0% (unpublished data). This is relatively low and more comparable with non-Asian series which feature overall ROMs of 35.6% and 31.3% (1,105) as compared to 70.5% and 60.4% in Asian series (1,103). This again suggests that the practice in Singapore is more closely aligned to Western practice, with higher rates of surgery as compared to active surveillance.

AUS-N vs. AUS-A: ROM

When AUS/FLUS was substratified into AUS-N *vs.* AUS-A, our pooled analysis showed that in both Asian and non-Asian cohorts, the ROM was significantly higher in AUS-N compared to AUS-A cases [49.6% *vs.* 17.0% (Asian) and 45.8% *vs.* 24.0% (non-Asian)]. This trend has also been demonstrated in the Singapore cohort, and independently demonstrated in many studies (5-13,16). This common finding between the two regions is strong support for the rationale of subclassifying the AUS/FLUS Bethesda category based on the presence of nuclear atypia.

Additionally, in the Singapore cohort, we also found that there was a higher rate of benign diagnoses on rFNA of AUS-A nodules than rFNA of AUS-N (70.6% *vs.* 48.7%, P=0.05), which is consistent with the findings of others, for example, those documented by Rosario *et al.* in their Brazilian series (9,89).

Within our institution, the terminology "AUS" is used for nodules with focal atypical nuclear features (nuclear enlargement, grooves, abnormal chromatin pattern, nuclear crowding and poorly formed inclusions), whereas "FLUS" refers to cases with some degree of architectural atypia (microfollicles, trabeculae, or crowding) without significant nuclear atypia. Although this is not a practice that is recommended in TBSRTC, we have found that others have also documented a similar terminology, namely in Brazil and Turkey (9,14,15), while yet others have adopted other terminology to denote subcategories (5,7,8).

Incidence of AUS-N vs. AUS-A

Interestingly, in the ten studies with adequate data for this comparison, we found stark geographical differences in the incidences of the subgroups. The incidence of AUS-N appeared to be far higher amongst Asian series compared to non-Asian cohorts (70.3% *vs.* 33.5%, P<0.001). It is noted that some Asian countries such as Japan may have somewhat narrower inclusion criteria for the AUS

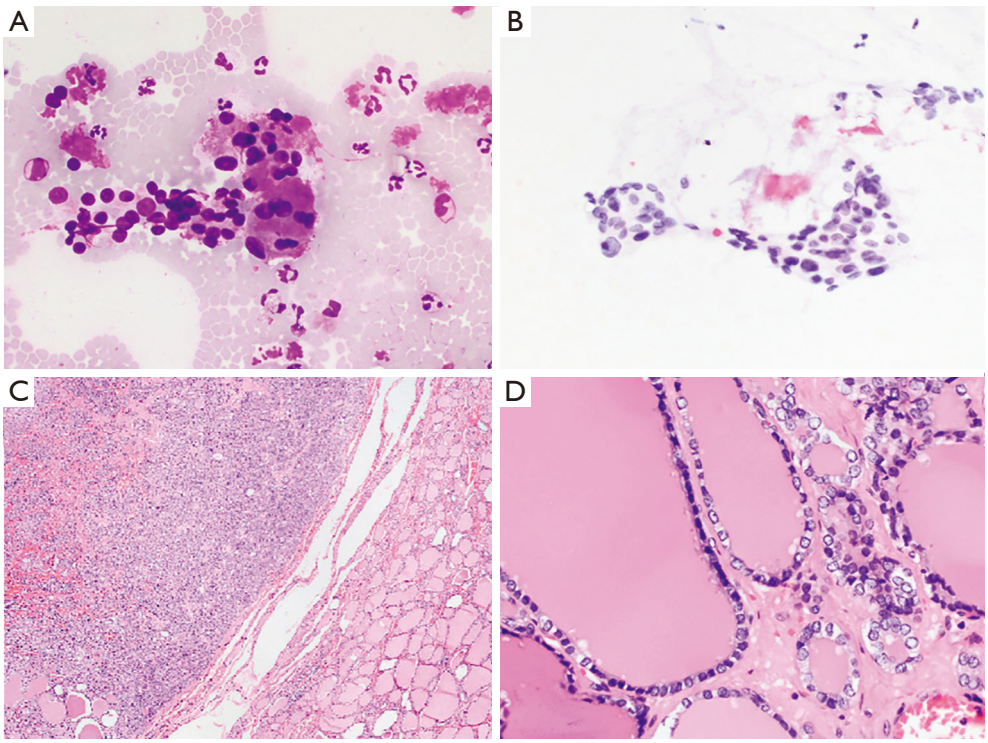

Figure 6 Example of AUS case with histologic resection. (A,B) This case was categorized as atypical (AUS) primarily because of nuclear features such as focal nuclear enlargement, occasional oval nuclei and variably pale chromatin (A: Diff-Quick stain, original magnification ×600; B: Papanicolaou stain, original magnification ×600). (C,D) Histology follow-up revealed a follicular adenoma (H&E stain, original magnification ×40, ×200). AUS, atypia of undetermined significance.

category than TBSRTC, primarily selecting for cases with PTC-like nuclear atypia (AUS-N), whilst cases with architectural atypia are categorized into the suspicious for FN/FN (SFN/FN) category instead (104). Although specific data from Japan was not available for the current analysis, our pooled incidences of studies from other Asian countries (Korea and Turkey) did show a higher incidence of AUS-N as opposed to non-Asian countries. Indeed, corroborating this, we also found a significantly lower incidence of FNs in resected AUS/FLUS nodules in Asian countries (mainland China, Taiwan, Korea and Turkey) compared to non-Asian countries (7.4% *vs.* 30.7%). These findings together may reflect the trend amongst Asian countries to reserve the use of AUS/FLUS to cases with PTC-like nuclear atypia, and the tendency to categorize cases with indeterminate architectural atypia into the SFN/FN category instead.

In view of the possible connection between cytomorphologic criteria of AUS-N and PTC, we examined the incidence of PTC in resected AUS/FLUS cases. Interestingly, we noticed that the Singapore cohort was an outlier with the lowest PTC incidence amongst the Asian studies (16.1%). After excluding the Singapore series, the PTC incidence in Asian *vs.* non-Asian studies was 50.2% *vs.* 29.1% respectively; P=0.009. Therefore, the relatively high incidence of PTC amongst resected cases may contribute to the higher prevalence of AUS-N cases in the Asian population. The low incidence of PTC in the Singapore AUS/FLUS cohort is intriguing, and raises the possibility of a difference in the true PTC disease prevalence of this population. Alternatively, this could potentially be accounted for by many PTC cases being cytologically classified into other Bethesda categories such as SM or Malignant. Another possible explanation is a high threshold for the histologic diagnosis of PTC, in particular, FVPTC, owing to the inherently subjective nature of interpretation of nuclear features. *Figure 6* illustrates a case which was classified as "AUS" due to the presence of some nuclear atypia, and histologically diagnosed as FA, although some degree of

nuclear pallor and enlargement was present.

Molecular practices

As we have shown in our results, there is a stark difference in the molecular approaches applied to AUS/FLUS nodules between Asian (chiefly mainland China and Korean) studies and those in the West (USA). Asian studies showed a much higher reliance on *BRAF* mutation testing (7,20,24,26-30), whilst Western studies favored broader multigene panel or array-based methods (74,81-84). None of the Asian or Western studies in the reviewed papers overlapped in their molecular interrogation methods.

We postulate that the preference for *BRAF* mutation testing in the Asian setting reflects the enrichment of PTCs in the AUS/FLUS category. This may be explained, at least in part, by the selection of cases with PTC-like nuclear atypia into this indeterminate category as discussed above. Furthermore, it appears that there is a higher rate of *BRAF* V600E mutation in Asian PTC cohorts than in Western cohorts, as illustrated in Korean centers where prevalence as high as 84% to 87% have been documented (106,107). This suggests that there is a greater proportion of classic PTC in Asian cohorts compared to FVPTC, given that the *BRAF* V600E mutation is more commonly detected in classical PTC. In support of this, the few studies that provided the breakdown of PTC variants showed that the percentage of FVPTCs amongst all PTCs was significantly lower in Asian series compared to non-Asian series (19.2% *vs.* 51.4%, P=0.006). This also bears out the point that *BRAF* mutations may be less helpful as a standalone test in non-Asian series, where not only is PTC less frequently encountered, but a greater proportion of the PTCs are FVPTCs, In Singapore, the *BRAF* mutation rate in PTC was 56% in a cohort from a single institution (108).

On the other side of the coin, *RAS* mutations feature more prominently in follicular-patterned neoplasms, which include FC, some FAs, NIFTPs and some FVPTCs. A loose comparison of the incidence of resected FNs within the AUS/FLUS category between Asian and non-Asian groups showed that these neoplasms featured more frequently in non-Asian than Asian cohorts (30.7% *vs.*7.4%, P<0.001). This ties in with our finding of higher AUS-A incidences in non-Asian cohorts compared to Asian series (57.6% *vs.* 22.3%, P<0.001). Altogether, this lends weight to the notion that Western practice as a whole have a greater tendency to classify cases with indeterminate architectural atypia as AUS/FLUS than the Asian practice. This may also partly explain why gene panel testing is the favored modality of molecular testing in the Western hemisphere.

Of course, cost and logistics also play a major role in the choice of molecular testing methods. Many of the commercially available gene panel tests and array-based tests were developed in the West, with limited accessibility in Asian countries.

In Singapore, molecular testing has not been validated in the local population and is not routinely performed in AUS/FLUS cases, however, individual tests may be performed on an ad-hoc basis. In such instances, the cost is considerable, as the aspirated material is usually transported to an accredited laboratory in the USA for testing, e.g., for the Thyroseq test. We have also found the results of repeat FNA to yield a more definitive cytologic diagnosis in 67.1% of cases, which is very helpful in the determination of the next management step.

Limitations

There are several limitations to this study. Firstly, it should be noted that the broad approach that we have adopted may have led to the inclusion of studies with selective reporting of some parameters. This may have introduced a further element of bias in addition to the inherent selection bias that accompanies retrospective studies, which makes up the majority of our data.

Secondly, there is a considerable amount of heterogeneity within each geographical cohort. Hence, the findings that we present here may be over generalized and not directly applicable to an individual institution's practices. This emphasizes the need for follow-up studies in individual practices, as there may be considerable inter- and intra-regional variations in the application of TBSRTC, disease prevalence and management considerations.

Thirdly, we detected an element of possible small study bias in the pooled analysis of several parameters, namely incidence, RR and OROM of AUS/FLUS, as well as the incidence of AUS-A and FN.

One caveat to our review of Singapore's practice is that it is drawn from a single tertiary institution's experience, and hence may not be fully representative of the country's practice as a whole. We also did not adjust for the impact of NIFTP on ROM as the number of cases in our series was too minimal for meaningful analysis (data not published). The impact of NIFTP has been comprehensively addressed by others and found to lower the ROM in indeterminate

nodules to a lesser degree in Asian cohorts than in the West (2,3,109-113).

Conclusions

This review provides insights into regional differences in the diagnosis and management of AUS/FLUS nodules between Asian and non-Asian countries, with a focused comparison of the local experience in Singapore. We found that although the overall incidence of AUS/FLUS was comparable, there were significant differences in the RRs and ROMs between Asian and non-Asian cohorts, which may reflect different collective thresholds for surgery. There remains much heterogeneity within each of the regional cohorts, as exemplified by the Singapore experience, where the RR and ROM appear to parallel the non-Asian studies despite having a largely Asian population. This strongly supports the recommendation of TBSRTC to validate ROM estimates in individual practices, particularly in this AUS/FLUS category that is most susceptible to subjective interpretation.

Additionally, we found that the subgroup of AUS-N showed a significantly higher ROM than AUS-A in both Asian and non-Asian series. These corroborating findings in both regions provides convincing grounds for subclassifying the AUS/FLUS Bethesda category based on the presence of nuclear atypia.

Lastly, Asian series yielded a higher incidence of PTC on surgical follow-up, while FVPTC and FNs featured more prominently in non-Asian series. This finding parallels the higher AUS-N incidence in Asia and the higher AUS-A incidence in non-Asian countries, suggesting variations in the application of TBSRTC criteria within the AUS/FLUS category. These differential incidences may influence and, indeed, explain the molecular approaches adopted in the different regions, as they inform us on the cost-effectiveness and predictive value considerations. For example, performing a single gene (*BRAF*) mutation test may be highly cost-effective in the Asian setting due to the higher proportion of cases with PTC-like nuclear atypia amongst AUS/FLUS cases. Hence, an awareness of regional variations in the incidences of specific histologic subtypes and the nuances in the interpretation of TBSRTC will help streamline choices as we collectively move toward greater accessibility to molecular testing in the indeterminate Bethesda categories.

Acknowledgments

Funding: None.

Footnote

Provenance and Peer Review: This article was commissioned by the Guest Editor (Kennichi Kakudo) for the series "Asian and Western Practice in Thyroid Pathology: Similarities and Differences" published in *Gland Surgery*. The article was sent for external peer review organized by the Guest Editor and the editorial office.

Peer Review File: Available at http://dx.doi.org/10.21037/gs-20-555

Conflicts of Interest: Both authors have completed the ICMJE uniform disclosure form (available at http://dx.doi.org/10.21037/gs-20-555). The series "Asian and Western Practice in Thyroid Pathology: Similarities and Differences" was commissioned by the editorial office without any funding or sponsorship. The authors have no other conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

References

1. Vuong HG, Ngo HTT, Bychkov A, et al. 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 2020;128:238-49.
2. Bychkov A, Keelawat S, Agarwal S, et al. 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;50:411-7.
3. Faquin WC, Wong LQ, Afrogheh AH, 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;124:181-7.
4. Cibas ES, Ali SZ. The 2017 Bethesda System for Reporting Thyroid Cytopathology. J Am Soc Cytopathol 2017;6:217-22.
5. Valderrabano P, Khazai L, Thompson ZJ, et al. Cancer Risk Stratification of Indeterminate Thyroid Nodules: A Cytological Approach. Thyroid 2017;27:1277-84.
6. VanderLaan PA, Marqusee E, Krane JF. Usefulness of Diagnostic Qualifiers for Thyroid Fine-Needle Aspirations With Atypia of Undetermined Significance. Am J Clin Pathol 2011;136:572-7.
7. Park HJ, Moon JH, Yom CK, et al. Thyroid 'atypia of undetermined significance' with nuclear atypia has high rates of malignancy and BRAF mutation. Cancer Cytopathol 2014;122:512-20.
8. Luu MH, Fischer AH, Stockl TJ, et al. Atypical follicular cells with equivocal features of papillary thyroid carcinoma is not a low-risk cytologic diagnosis. Acta Cytol 2011;55:526-30.
9. Rosario PW. Thyroid nodules with atypia or follicular lesions of undetermined significance (Bethesda Category III): importance of ultrasonography and cytological subcategory. Thyroid 2014;24:1115-20.
10. Pagni F, Prada M, Goffredo P, et al. 'Indeterminate for malignancy' (Tir3/Thy3 in the Italian and British systems for classification) thyroid fine needle aspiration (FNA) cytology reporting: morphological criteria and clinical impact. Cytopathology 2014;25:170-6.
11. Mosca L, Silva LFFD, Carneiro PC, et al. Malignancy rates for Bethesda III subcategories in thyroid fine needle aspiration biopsy (FNAB). Clinics (Sao Paulo) 2018;73:e370.
12. Wu HH, Inman A, Cramer HM. Subclassification of "atypia of undetermined significance" in thyroid fine-needle aspirates: Atypia of Undetermined Significance. Diagn Cytopathol 2014;42:23-9.
13. Onder S, Firat P, Ates D. The Bethesda system for reporting thyroid cytopathology: an institutional experience of the outcome of indeterminate categories. Cytopathology 2014;25:177-84.
14. Çuhaci N, Arpaci D, Üçler R, et al. Malignancy rate of thyroid nodules defined as follicular lesion of undetermined significance and atypia of undetermined significance in thyroid cytopathology and its relation with ultrasonographic features. Endocr Pathol 2014;25:248-56.
15. Dincer N, Balci S, Yazgan A, et al. Follow-up of atypia and follicular lesions of undetermined significance in thyroid fine needle aspiration cytology. Cytopathology 2013;24:385-90.
16. Gan TRX, Nga ME, Lum JHY, et al. Thyroid cytology-nuclear versus architectural atypia within the 'Atypia of undetermined significance/follicular lesion of undetermined significance' Bethesda category have significantly different rates of malignancy. Cancer Cytopathol 2017;125:245-56.
17. Keelawat S, Rangdaeng S, Koonmee S, et al. Current Status of Thyroid Fine-Needle Aspiration Practice in Thailand. J Pathol Transl Med 2017;51:565-70.
18. Thewjitcharoen Y, Butadej S, Nakasatien S, et al. Incidence and malignancy rates classified by The Bethesda System for Reporting Thyroid Cytopathology (TBSRTC) - An 8-year tertiary center experience in Thailand. J Clin Transl Endocrinol 2018;16:100175.
19. Mao F, Xu HX, Zhao CK, et al. Thyroid imaging reporting and data system in assessment of cytological Bethesda Category III thyroid nodules. Clin Hemorheol Microcirc 2017;65:163-73.
20. Zheng B, Zarka MA, Chen C, et al. The largest CAP-certified Chinese reference laboratory experience with the Bethesda system for reporting thyroid cytopathology: correlation with histologic and BRAF data. J Am Soc Cytopathol 2018;7:16-21.
21. Ke J, Jianyong L, Ying L, et al. The use of The Bethesda System for Reporting Thyroid Cytopathology in a Chinese population: An analysis of 13 351 specimens. Diagn Cytopathol 2019;47:876-80.
22. Jan IS, Lee YT, Wang CM, et al. The surgery and repeat aspiration outcomes of the atypia of undetermined significance/follicular lesion of undetermined significance category in The Bethesda System for Reporting Thyroid Cytopathology. Asian J Surg 2019;42:144-7.
23. Satoh S, Yamashita H, Kakudo K. Thyroid Cytology: The Japanese System and Experience at Yamashita Thyroid Hospital. J Pathol Transl Med 2017;51:548-54.
24. Kim SK, Hwang TS, Yoo YB, et al. Surgical Results of Thyroid Nodules according to a Management Guideline Based on the BRAF V600E Mutation Status. J Clin Endocrinol Metab 2011;96:658-64.

25. Chung YS, Yoo C, Jung JH, et al. Review of atypical cytology of thyroid nodule according to the Bethesda system and its beneficial effect in the surgical treatment of papillary carcinoma. J Korean Surg Soc 2011;81:75-84.
26. Hyeon J, Ahn S, Shin JH, et al. The prediction of malignant risk in the category 'atypia of undetermined significance/follicular lesion of undetermined significance' of the Bethesda System for Reporting Thyroid Cytopathology using subcategorization and BRAF mutation results. Cancer Cytopathol 2014;122:368-76.
27. Yoo MR, Gweon HM, Park AY, et al. Repeat Diagnoses of Bethesda Category III Thyroid Nodules: What To Do Next? PloS One 2015;10:e0130138.
28. Jung YY, Jung S, Lee HW, et al. Significance of Subcategory Atypia of Undetermined Significance/ Follicular Lesion of Undetermined Significance Showing Both Cytologic and Architectural Atypia in Thyroid Aspiration Cytology. Acta Cytol 2015;59:370-6.
29. Koh J, Moon HJ, Kim EK, et al. The 5-tiered categorization system for reporting cytology is sufficient for management of patients with thyroid nodules compared to the 6-tiered Bethesda system. Endocrine 2016;53:489-96.
30. Kim SJ, Roh J, Baek JH, et al. Risk of malignancy according to sub-classification of the atypia of undetermined significance or follicular lesion of undetermined significance (AUS/FLUS) category in the Bethesda system for reporting thyroid cytopathology. Cytopathology 2017;28:65-73.
31. Kim SD, Han SH, Jeong WJ, et al. Differences in Clinical Features Between Subcategories of 'Atypia/Follicular Lesion of Undetermined Significance'. Endocr Pathol 2017;28:247-52.
32. Kim M, Park HJ, Min HS, et al. The Use of the Bethesda System for Reporting Thyroid Cytopathology in Korea: A Nationwide Multicenter Survey by the Korean Society of Endocrine Pathologists. J Pathol Transl Med 2017;51:410-7.
33. Hong SH, Lee H, Cho MS, et al. Malignancy Risk and Related Factors of Atypia of Undetermined Significance/ Follicular Lesion of Undetermined Significance in Thyroid Fine Needle Aspiration. Int J Endocrinol 2018;2018:4521984.
34. Mondal SK, Sinha S, Basak B, et al. The Bethesda system for reporting thyroid fine needle aspirates: A cytologic study with histologic follow-up. J Cytol 2013;30:94-9.
35. Mehra P, Verma AK. Thyroid cytopathology reporting by the bethesda system: a two-year prospective study in an academic institution. Patholog Res Int 2015;2015:240505.
36. Garg S, Desai NJ, Mehta D, et al. To Establish Bethesda System for Diagnosis of Thyroid Nodules on the Basis of Fnac with Histopathological Correlation. J Clin Diagn Res 2015;9:EC17-21.
37. Arul P, Akshatha C, Masilamani S. A study of malignancy rates in different diagnostic categories of the Bethesda system for reporting thyroid cytopathology: An institutional experience. Biomed J 2015;38:517-22.
38. Mahajan S, Srinivasan R, Rajwanshi A, et al. Risk of Malignancy and Risk of Neoplasia in the Bethesda Indeterminate Categories: Study on 4,532 Thyroid Fine-Needle Aspirations from a Single Institution in India. Acta Cytol 2017;61:103-10.
39. Kumari KA, Jadhav PD, Prasad C, et al. Diagnostic Efficacy of Ultrasound-Guided Fine Needle Aspiration Combined with the Bethesda System of Reporting. J Cytol 2019;36:101-5.
40. Naz S, Hashmi AA, Khurshid A, et al. Diagnostic accuracy of Bethesda system for reporting thyroid cytopathology: an institutional perspective. Int Arch Med 2014;7:46.
41. Ozluk Y, Pehlivan E, Gulluoglu MG, et al. The use of the Bethesda terminology in thyroid fine-needle aspiration results in a lower rate of surgery for nonmalignant nodules: a report from a reference center in Turkey. Int J Surg Pathol 2011;19:761-71.
42. Ustün H, Astarcı HM, Altunkaya C, et al. Fine-needle aspiration of follicular patterned lesions of the thyroid: diagnosis, management, and follow-up according to thyroid Bethesda system. Acta Cytol 2012;56:361-9.
43. Firat P, Cochand-Priollet B. The Bethesda system for reporting thyroid fine needle aspiration cytology: a study comparing the results of two centers from two different countries. Ann Pathol 2012;32:e29-34, 415-20.
44. Tepeoğlu M, Bilezikçi B, Bayraktar SG. A histological assessment of the Bethesda system for reporting thyroid cytopathology (2010) abnormal categories: a series of 219 consecutive cases. Cytopathology 2014;25:39-44.
45. Muratli A, Erdogan N, Sevim S, et al. Diagnostic efficacy and importance of fine-needle aspiration cytology of thyroid nodules. J Cytol 2014;31:73-8.
46. Gocun PU, Karakus E, Bulutay P, et al. What is the malignancy risk for atypia of undetermined significance? Three years' experience at a university hospital in Turkey. Cancer Cytopathol 2014;122:604-10.
47. Kuru B, Atmaca A, Kefeli M. Malignancy rate associated with Bethesda category III (AUS/FLUS) with and without repeat fine needle aspiration biopsy. Diagn Cytopathol

2016;44:394-8.
48. Selek A, Cetinarslan B, Kıvrakoğlu E, et al. Histologic outcome of thyroid nodules with repeated diagnosis of atypia in thyroid fine-needle aspiration biopsy. Future Oncol 2016;12:801-5.
49. Turkyilmaz S, Ulusahin M, Celebi B, et al. Thyroid nodules classified as atypia or follicular lesions of undetermined significance deserve further research: Analysis of 305 surgically confirmed nodules. Cytopathology 2017;28:391-9.
50. Öcal B, Korkmaz MH, Yılmazer D, et al. The Malignancy Risk Assessment of Cytologically Indeterminate Thyroid Nodules Improves Markedly by Using a Predictive Model. Eur Thyroid J 2019;8:83-9.
51. Mufti ST, Molah R. The bethesda system for reporting thyroid cytopathology: a five-year retrospective review of one center experience. Int J Health Sci (Qassim) 2012;6:159-73.
52. Al-Abbadi MA, Shareef SQ, Ali JA, et al. Application of the Bethesda System for Reporting Thyroid Cytopathology in the Eastern Province of Saudi Arabia: phase I pilot retrospective analysis. Acta Cytol 2013;57:481-8.
53. Alabdulqader NA, Shareef SQ, Ali JA, et al. Application of the Bethesda System for Reporting Thyroid Cytopathology in the Eastern Province of Saudi Arabia: A Follow-Up Study. Acta Cytol 2015;59:233-8.
54. Al Dawish MA, Robert AA, Muna A, et al. Bethesda System for Reporting Thyroid Cytopathology: A three-year study at a tertiary care referral center in Saudi Arabia. World J Clin Oncol 2017;8:151-7.
55. Hirsch D, Robenshtok E, Bachar G, et al. The Implementation of the Bethesda System for Reporting Thyroid Cytopathology Improves Malignancy Detection Despite Lower Rate of Thyroidectomy in Indeterminate Nodules. World J Surg 2015;39:1959-65.
56. Ronen O, Cohen H, Abu M. Review of a single institution's fine needle aspiration results for thyroid nodules: Initial observations and lessons for the future. Cytopathology 2019;30:468-74.
57. Sinna EA, Ezzat N. Diagnostic accuracy of fine needle aspiration cytology in thyroid lesions. J Egypt Natl Canc Inst 2012;24:63-70.
58. Al-Abbadi MA, Shareef SQ, Yousef MM, et al. A follow-up study on thyroid aspirates reported as atypia of undetermined significance/follicular lesion of undetermined significance and follicular neoplasm/suspicious for follicular neoplasm: A multicenter study from the Arabian Gulf region. Diagn Cytopathol 2017;45:983-8.
59. Jo VY, Stelow EB, Dustin SM, et al. Malignancy risk for fine-needle aspiration of thyroid lesions according to the Bethesda System for Reporting Thyroid Cytopathology. Am J Clin Pathol 2010;134:450-6.
60. Rabaglia JL, Kabbani W, Wallace L, et al. Effect of the Bethesda system for reporting thyroid cytopathology on thyroidectomy rates and malignancy risk in cytologically indeterminate lesions. Surgery 2010;148:1267-72; discussion 1272-3.
61. Olson MT, Clark DP, Erozan YS, et al. Spectrum of risk of malignancy in subcategories of 'atypia of undetermined significance'. Acta Cytol 2011;55:518-25.
62. Bongiovanni M, Crippa S, Baloch Z, et al. Comparison of 5-tiered and 6-tiered diagnostic systems for the reporting of thyroid cytopathology. Cancer Cytopathol 2012;120:117-25.
63. Chen JC, Pace SC, Chen BA, et al. Yield of repeat fine-needle aspiration biopsy and rate of malignancy in patients with atypia or follicular lesion of undetermined significance: the impact of the Bethesda System for Reporting Thyroid Cytopathology. Surgery 2012;152:1037-44.
64. Horne MJ, Chhieng DC, Theoharis C, et al. Thyroid follicular lesion of undetermined significance: Evaluation of the risk of malignancy using the two-tier sub-classification. Diagn Cytopathol 2012;40:410-5.
65. Harvey AM, Mody DR, Amrikachi M. Thyroid fine-needle aspiration reporting rates and outcomes before and after Bethesda implementation within a combined academic and community hospital system. Arch Pathol Lab Med 2013;137:1664-8.
66. Nagarkatti SS, Faquin WC, Lubitz CC, et al. Management of thyroid nodules with atypical cytology on fine-needle aspiration biopsy. Ann Surg Oncol 2013;20:60-5.
67. Olson MT, Boonyaarunnate T, Aragon Han P, et al. A tertiary center's experience with second review of 3885 thyroid cytopathology specimens. J Clin Endocrinol Metab 2013;98:1450-7.
68. Theoharis C, Adeniran AJ, Roman S, et al. The impact of implementing The Bethesda System for reporting of thyroid FNA at an academic center. Diagn Cytopathol 2013;41:858-63.
69. Broome JT, Cate F, Solorzano CC. Utilization and impact of repeat biopsy for follicular lesion/atypia of undetermined significance. World J Surg 2014;38:628-33.
70. Ho AS, Sarti EE, Jain KS, et al. Malignancy rate in thyroid

nodules classified as Bethesda category III (AUS/FLUS). Thyroid 2014;24:832-9.
71. Lee B, Smola B, Roh MH, et al. The impact of using the Bethesda System for reporting thyroid cytology diagnostic criteria on the follicular lesion of undetermined significance category. J Am Soc Cytopathol 2014;3:131-6.
72. Mathur A, Najafian A, Schneider EB, et al. Malignancy risk and reproducibility associated with atypia of undetermined significance on thyroid cytology. Surgery 2014;156:1471-6.
73. McElroy MK, Mahooti S, Hasteh F. A single institution experience with the new bethesda system for reporting thyroid cytopathology: correlation with existing cytologic, clinical, and histological data. Diagn Cytopathol 2014;42:564-9.
74. Sullivan PS, Hirschowitz SL, Fung PC, et al. The impact of atypia/follicular lesion of undetermined significance and repeat fine-needle aspiration: 5 years before and after implementation of the Bethesda System. Cancer Cytopathol 2014;122:866-72.
75. Walts AE, Mirocha J, Bose S. Follicular lesion of undetermined significance in thyroid FNA revisited. Diagn Cytopathol 2014;42:18-22.
76. Deniwar A, Hambleton C, Thethi T, et al. Examining the Bethesda criteria risk stratification of thyroid nodules. Pathol Res Pract 2015;211:345-8.
77. Kantola S, Virani N, Haus C, et al. Prospective evaluation of impact of using the Bethesda System for Reporting Thyroid Cytopathology: an institutional experience. J Am Soc Cytopathol 2015;4:25-9.
78. Brandler TC, Aziz MS, Coutsouvelis C, et al. Young investigator challenge: Atypia of undetermined significance in thyroid FNA: Standardized terminology without standardized management--a closer look at repeat FNA and quality measures. Cancer Cytopathol 2016;124:37-43.
79. Krauss EA, Mahon M, Fede JM, et al. Application of the Bethesda Classification for Thyroid Fine-Needle Aspiration: Institutional Experience and Meta-analysis. Arch Pathol Lab Med 2016;140:1121-31.
80. Shrestha RT, Hennessey JV. Cytologic subclassification of atypia of undetermined significance may predict thyroid nodules more likely to be malignant at surgery. Diagn Cytopathol 2016;44:492-8.
81. Valderrabano P, Leon ME, Centeno BA, et al. Institutional prevalence of malignancy of indeterminate thyroid cytology is necessary but insufficient to accurately interpret molecular marker tests. Eur J Endocrinol 2016;174:621-9.
82. Guo A, Kaminoh Y, Forward T, et al. Fine Needle Aspiration of Thyroid Nodules Using the Bethesda System for Reporting Thyroid Cytopathology: An Institutional Experience in a Rural Setting. Int J Endocrinol 2017;2017:9601735.
83. Deaver KE, Haugen BR, Pozdeyev N, et al. Outcomes of Bethesda categories III and IV thyroid nodules over 5 years and performance of the Afirma gene expression classifier: A single-institution study. Clin Endocrinol (Oxf) 2018;89:226-32.
84. Seagrove-Guffey MA, Hatic H, Peng H, et al. Malignancy rate of atypia of undetermined significance/follicular lesion of undetermined significance in thyroid nodules undergoing FNA in a suburban endocrinology practice: A retrospective cohort analysis. Cancer Cytopathol 2018;126:881-8.
85. Bresler A, Mehta V, Schiff BA, et al. Comparison of Bethesda cytopathology classification to surgical pathology across racial-ethnic groups. Head Neck 2019;41:2340-5.
86. Williams BA, Bullock MJ, Trites JR, et al. Rates of thyroid malignancy by FNA diagnostic category. J Otolaryngol Head Neck Surg 2013;42:61.
87. Bernstein JM, Shah M, MacMillan C, et al. Institution-specific risk of papillary thyroid carcinoma in atypia/follicular lesion of undetermined significance: Risk of PTC in Atypia/Follicular Lesion of Undetermined Significance. Head Neck 2016;38:E1210-5.
88. Erivwo P, Ghosh C. Atypia of Undetermined Significance in Thyroid Fine-Needle Aspirations: Follow-Up and Outcome Experience in Newfoundland, Canada. Acta Cytol 2018;62:85-92.
89. Rosario PW, Calsolari MR. Importance of cytological subclassification of thyroid nodules with Bethesda category III cytology (AUS/FLUS) into architectural atypia only and nuclear atypia: A prospective study. Diagn Cytopathol 2017;45:604-7.
90. Reuters KB, Mamone MCOC, Ikejiri ES, et al. Bethesda Classification and Cytohistological Correlation of Thyroid Nodules in a Brazilian Thyroid Disease Center. Eur Thyroid J 2018;7:133-8.
91. Ratour J, Polivka M, Dahan H, et al. Diagnosis of follicular lesions of undetermined significance in fine-needle aspirations of thyroid nodules. J Thyroid Res 2013;2013:250347.
92. Stanek-Widera A, Biskup-Frużyńska M, Zembala-

Nożyńska E, et al. Clinical importance of follicular lesion of undetermined significance (diagnostic category III according to Bethesda System) diagnosed from Fine-Needle Aspiration Biopsy. Endokrynol Pol 2016;67:12-6.
93. Estrada Muñoz L, Díaz Del Arco C, Ortega Medina L, et al. Thyroid Atypia/Follicular Lesion of Undetermined Significance: Attitudes towards the Diagnosis of Bethesda System III Nodules. Acta Cytol 2017;61:21-6.
94. Paajanen I, Metso S, Jaatinen P, et al. Thyroid FNA diagnostics in a real-life setting: Experiences of the implementation of the Bethesda system in Finland. Cytopathology 2018;29:189-95.
95. Mileva M, Stoilovska B, Jovanovska A, et al. Thyroid cancer detection rate and associated risk factors in patients with thyroid nodules classified as Bethesda category III. Radiol Oncol 2018;52:370-6.
96. Sarkis LM, Norlen O, Aniss A, et al. The Australian experience with the Bethesda classification system for thyroid fine needle aspiration biopsies. Pathology 2014;46:592-5.
97. Fatman L, Michelow P. Thyroid cytopathology with an emphasis on the 'atypical cells of uncertain significance' category: a 3-year audit with cytohistologic correlation. Acta Cytol 2015;59:17-25.
98. Ryu YJ, Jung YS, Yoon HC, et al. Atypia of undetermined significance on thyroid fine needle aspiration: surgical outcome and risk factors for malignancy. Ann Surg Treat Res 2014;86:109-14.
99. Srbova L, Gabalec F, Ryska A, et al. Results of retrospective classification of thyroid FNAs according to the Bethesda system: would this have improved accuracy? Cytopathology 2015;26:231-7.
100. Rossi M, Lupo S, Rossi R, et al. Proposal for a novel management of indeterminate thyroid nodules on the basis of cytopathological subclasses. Endocrine 2017;57:98-107.
101. Labarge B, Walter V, Lengerich EJ, et al. Evidence of a positive association between malpractice climate and thyroid cancer incidence in the United States. PLoS One 2018;13:e0199862.
102. Kakudo K, Bychkov A, Abelardo A, et al. Malpractice Climate Is a Key Difference in Thyroid Pathology Practice Between North America and the Rest of the World. Arch Pathol Lab Med 2019;143:1171.
103. Kakudo K, Higuchi M, Hirokawa M, et al. Thyroid FNA cytology in Asian practice-Active surveillance for indeterminate thyroid nodules reduces overtreatment of thyroid carcinomas. Cytopathology 2017;28:455-66.
104. Kakudo K, Kameyama K, Miyauchi A, et al. Introducing the reporting system for thyroid fine-needle aspiration cytology according to the new guidelines of the Japan Thyroid Association. Endocr J 2014;61:539-52.
105. Bongiovanni M, Spitale A, Faquin WC, et al. The Bethesda System for Reporting Thyroid Cytopathology: a meta-analysis. Acta Cytol 2012;56:333-9.
106. Kim SK, Kim DL, Han HS, et al. Pyrosequencing analysis for detection of a BRAFV600E mutation in an FNAB specimen of thyroid nodules. Diagn Mol Pathol 2008;17:118-25.
107. Kim SK, Song KH, Lim SD, et al. Clinical and pathological features and the BRAF(V600E) mutation in patients with papillary thyroid carcinoma with and without concurrent Hashimoto thyroiditis. Thyroid 2009;19:137-41.
108. Goh X, Lum J, Yang SP, et al. BRAF mutation in papillary thyroid cancer-Prevalence and clinical correlation in a South-East Asian cohort. Clin Otolaryngol 2019;44:114-23.
109. Bongiovanni M, Faquin W, Giovanella L, et al. 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. Eur J Endocrinol 2019;181:389-96.
110. Zhou H, Baloch ZW, Nayar R, 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;126:20-6.
111. Ruanpeng D, Cheungpasitporn W, Thongprayoon C, et al. Systematic review and meta-analysis of the impact of Noninvasive Follicular Thyroid Neoplasm with Papillary-Like Nuclear Features (NIFTP) on cytological diagnosis and thyroid cancer prevalence. Endocr Pathol 2019;30:189-200.
112. Higuchi M, Hirokawa M, Kanematsu R, et al. Impact of the modification of the diagnostic criteria in the 2017 Bethesda System for Reporting Thyroid Cytopathology: a report of a single institution in Japan. Endocr J 2018;65:1193-8.
113. Vuong HG, Tran TTK, Bychkov A, et al. 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;25:491-502.

Cite this article as: Ooi LY, Nga ME. Atypia of undetermined significance/follicular lesion of undetermined significance: Asian *vs.* non-Asian practice, and the Singapore experience. Gland Surg 2020;9(5):1764-1787. doi: 10.21037/gs-20-555

Supplementary

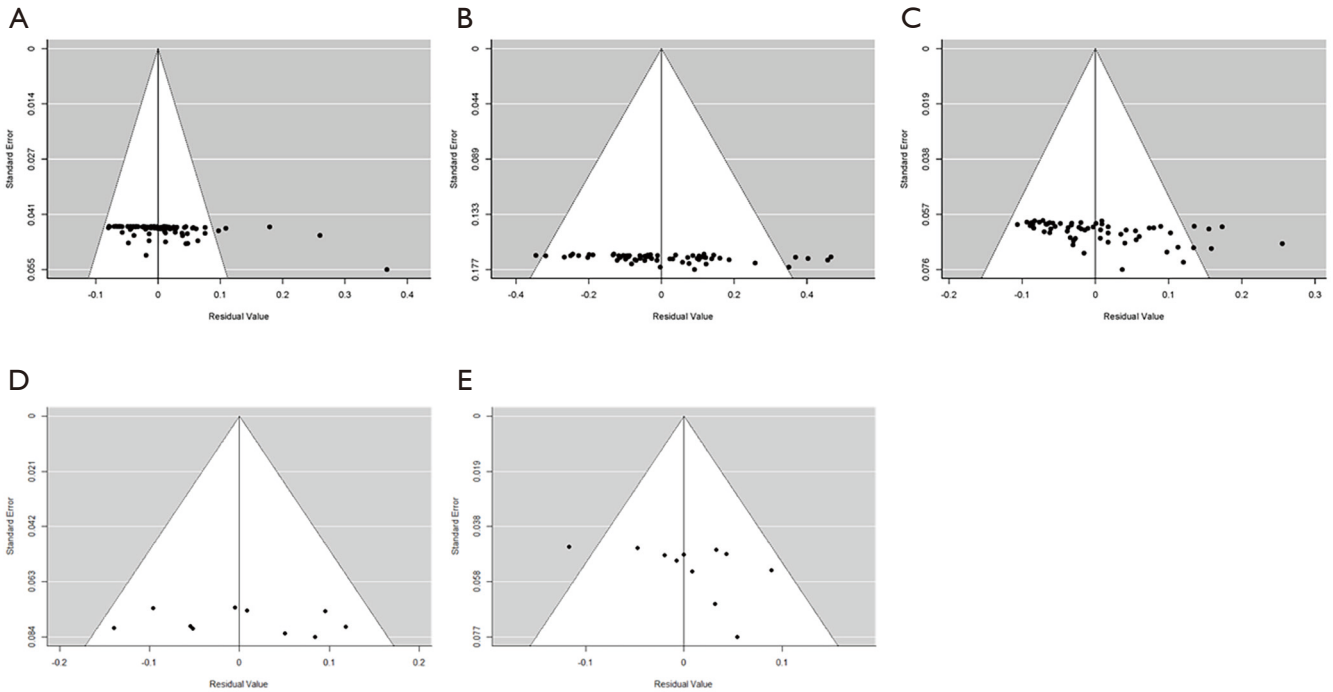

Figure S1 Funnel plots of sample estimates from studies included in pooled analysis of the following: (A) incidence of AUS/FLUS; (B) RR of AUS/FLUS; (C) OROM of AUS/FLUS; (D) incidence of AUS-A; (E) incidence of FN. AUS/FLUS, atypia of undetermined significance/follicular lesion of undetermined significance; RR, resection rate; OROM, overall risk of malignancy; AUS-A, AUS/FLUS with architectural atypia; FN, follicular neoplasm.

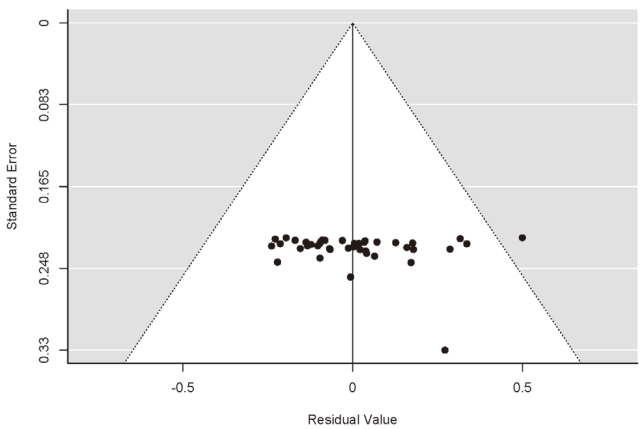

Figure S2 Funnel plot of sample estimates from studies that compared ROMs between the cytomorphological subgroups (AUS-A *vs.* AUS-N). AUS-A, AUS/FLUS with architectural atypia; AUS-N, AUS/FLUS with nuclear atypia; ROM, risk of malignancy.