S Afr J Surg. 2026;64:12-16. Online first
https://doi.org/10.36303/SAJS.02294

S Afr J Surg
ISSN: 0038-2361 eISSN: 2078-5151

ENDOCRINE SURGERY

The performance of fine needle aspiration cytology (FNAC) compared to histology in the preoperative diagnosis of thyroid cancer: a single centre study

S Pillay,[1] T Ntrama,[2] G Muthen,[3] C Trainer,[4] Y Varachia,[5] M Mmutlane,[6] T Luvhengo,[7] N Sishuba[7]

[1] Department of Paediatrics, Chris Hani Baragwaneth Academic Hospital, South Africa
[2] Potchefstroom Hospital, South Africa
[3] Department of Obstetrics and Gynaecology, Chris Hani Baragwaneth Academic Hospital, South Africa
[4] Department of Internal Medicine, Harry Gwala Regional Hospital, South Africa
[5] Department of Internal Medicine, Chris Hani Baragwaneth Academic Hospital, South Africa
[6] Department of Internal Medicine, Leratong Hospital, South Africa
[7] Department of Surgery, Charlotte Maxeke Johannesburg Academic Hospital, South Africa

Corresponding author, email: *pillaysudarshan@gmail.com*

Background: Thyroid cancer is the most common endocrine malignancy globally. Fine needle aspiration cytology (FNAC) is used in the diagnostic workup of thyroid nodules. There have been discrepancies between FNAC and histology at Charlotte Maxeke Johannesburg Academic Hospital (CMJAH). This prompted a study to determine the performance of FNAC compared to histology in diagnosing thyroid cancer at CMJAH.
Methods: A retrospective review of patients who underwent thyroid surgery at CMJAH between 1 January 2015 and 1 September 2022 was conducted. Patient details were accessed via the theatre registry and the department of nuclear medicine. FNAC and histology results were accessed using the National Health Laboratory Service (NHLS) website. The data was collected and entered into a Microsoft excel spreadsheet. STATA statistical software was used to analyse the data.
Results: A total of 369 patients were screened, of which 222 were included in the study. We found the prevalence of thyroid cancer at CMJAH to be 35.6%. FNAC was found to have a sensitivity of 83.6% and a specificity of 88.7%. Its accuracy was 0.861 and it had a PPV of 81.2% and a NPV of 90.3%. The likelihood ratio generated was 7.39. PTC was found to be the most prevalent subtype with the FNAC malignancy detection rate of 85%.
Conclusion: Our study showed that FNAC performed well in diagnosing thyroid cancer preoperatively. Therefore, it should continue to be used in the diagnostic workup of thyroid nodules at CMJAH.
Keywords: thyroid cancer, fine needle aspiration cytology, diagnostic, accuracy, sensitivity, specificity

Introduction

Fine needle aspiration cytology (FNAC) is a minimally invasive and inexpensive diagnostic technique used to confirm whether a lesion is benign or malignant.[1] FNACs are associated with very few complications and are regarded as the first-line diagnostic technique in examining thyroid nodules.[1] This diagnostic tool may be used to prevent unnecessary surgery and is used to determine a surgical approach if malignancy of the thyroid is suspected.

The initial workup for any newly diagnosed thyroid nodule should include a serum thyroid-stimulating hormone (TSH) level. Measurement of TSH allows differentiation between functional and non-functional nodules. Ultrasound is the first imaging modality for non-functional nodules, to determine if the nodule has features suspicious of malignancy. Imaging is then followed by FNAC of any thyroid nodule ≥ 1 cm with suspicious features (i.e., irregular borders, taller than wide nodules, central vascularity, loss of halo sign, and presence of micro-calcifications).[2] FNAC can be performed palpation-guided (PGFNA) and ultrasound-guided (USGFNA) with or without rapid on-site evaluation (ROSE). USGFNA with ROSE requires the cytopathologist to prepare the smears and analyse them as soon as the thyroid nodule is aspirated.[1] The ROSE method minimises the risk of non-diagnostic or unsatisfactory specimens. FNAC results are universally reported using the Bethesda System for reporting thyroid cytopathology. This reporting system is standardised and grades FNAC results into six categories, namely (1) non-diagnostic or unsatisfactory, (2) benign, (3) atypia of undetermined significance (AUS) or follicular lesion of undetermined significance (FLUS), (4) follicular neoplasm or suspicious for a follicular neoplasm, (5) suspicious for malignancy, and (6) malignant. A malignancy risk is assigned to each category ranging from 1% in category 1 to 99% in category 6.[1]

The accuracy of FNAC compared with postoperative histology results in the diagnosis of thyroid cancer has previously been studied.[3]

FNAC has the potential to offer great benefit in the South African (SA) setting. It could indicate when surgery is or is not necessary, thus sparing the use of needless resources and complications. A local study done by Conradie et al.[3] in 2024 evaluated the performance of FNAC in low- and middle-income countries. It was a national study, which included two private and one public laboratories. They found that FNAC was non-diagnostic in 38% of the cases. It had a diagnostic accuracy of 74%, sensitivity of 73%, specificity of 74%, positive predictive value (PPV) of 67% and negative predictive value (NPV) of 79%. The authors concluded that the diagnostic accuracy of FNAC needs further investigation in the different institutions in SA, motivating the decision to conduct this study.

This study primarily aimed to investigate the diagnostic performance and overall accuracy of FNAC in diagnosing thyroid cancer at CMJAH. The secondary aims were to determine the relation of FNAC to the size of the thyroid nodule and thyroid cancer subtype, and the prevalence of thyroid cancer in our setting.

Patients and methods

This was a retrospective cross-sectional study conducted at the endocrine surgery unit of the department of surgery at CMJAH. All the patients who underwent thyroid surgery, irrespective of their indication for surgery, following FNAC at CMJAH between 1 January 2015 and 1 September 2022 were included in the study. There was a standardised approach to the performance of the FNACs. They were done by the same team, i.e., interventional radiology team, using the ultrasound guided technique with the cytopathologist on site to do ROSE. Patients under 18 years of age, those who had previous thyroid surgery, recurrent thyroid cancer, incomplete clinical records, or with no FNAC results were excluded.

The theatre registry and the department of nuclear medicine at CMJAH were interrogated for all the patients who had undergone thyroid surgeries within our study period. From this we collected age, sex and type of surgery performed. The preoperative FNAC and the postoperative histology results were retrospectively accessed via the NHLS online server. When the nodule size was unavailable from the NHLS, we used the Picture Archiving and Communication System (PACS) which is a radiology and imaging reporting database.

Statistical analysis

The data was collected and entered into a Microsoft Excel spreadsheet where it was de-identified to maintain patient confidentiality. It was then analysed using the STATA statistical software. We compared the FNAC results to the histology results to check if they were concordant. Based on this comparison, we were able to classify the FNAC result into a true positive (TP), true negative (TN), false positive (FP), or false negative (FN) result. We then calculated the sensitivity, specificity, accuracy, PPV, NPV of the FNAC using the calculations which are given below. We used tables to display the numerical data, e.g., age and figures to display the categorical variables, e.g., sex, FNAC results, etc.

- Formulae:
  Accuracy = TP+TN/TP+FP+TN+FN
  Sensitivity = TP/TP+FN
  Specificity = TN/TN+FP

Results

Over the study period, a total of 369 patients who had thyroid surgery were identified from both the theatre registry and department of nuclear medicine. After applying our exclusion criteria, 222 patients were included in the study. The median age was 45–54 years and the majority of patients were > 45 years of age (Table I).

In this study, 86% of patients were female compared to 14% who were male.

Out of the 222 patients we included in our study, 79 of the patients had confirmed thyroid cancer by histology, therefore, thyroid cancers have a prevalence of 35.6% at CMJAH out of the selected population.

Of all surgeries done, 94% were total thyroidectomies and only 6% were lobectomies.

Table I: Age groups of the study participants

Age groups	Proportion of study participants (%)
18–24	3.6
25–34	14.4
35–44	18.5
45–54	32
>/= 55	31.5
Total	100

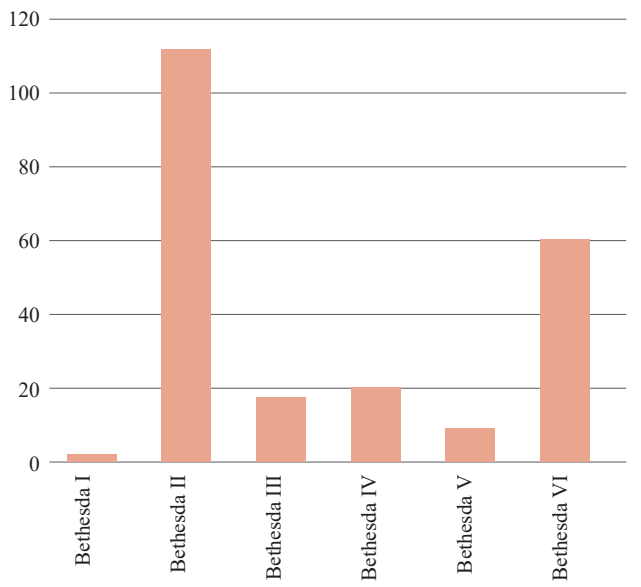

Figure 1: Proportion of all Bethesda categories

			[95% Confidence Interval]	
Prevalence	Pr(A)	37%	30%	44.3%
Sensitivity	Pr(+|A)	83.6%	72.5%	91.5%
Specificity	Pr(-|N)	88.7%	81.4%	93.8%
ROC area	(Sens. + Spec.)/2	.861	.808	.915
Likelihood ratio (+)	Pr(+|A)/Pr(+|N)	7.39	4.38	12.5
Likelihood ratio (-)	Pr(-|A)/Pr(-|N)	.185	.107	.319
Odds ratio	LR(+)/LR(-)	39.9	16.9	94.4
Positive predictive value	Pr(A|+)	81.2%	69.9%	89.6%
Negative predictive value	Pr(N|-)	90.3%	83.2%	95%

Figure 2: Sensitivity, specificity, accuracy, PPV, and NPV of Bethesda II, V, and VI – and thus the diagnostic performance of FNAC

Table II: TP, TN, FP, FN of Bethesda II, V and VI

	Malignant (histology)	Benign (histology)	Total
FNAC+ (positive result)	56 true positives	13 false positives	69
FNAC- (negative result)	11 false negatives	102 true negatives	113
Total	67	115	182

Table III: TP, TN, FP, FN, of Bethesda II, V & VI

	Bethesda II				Bethesda V				Bethesda VI		
	Malignant	Benign	Total		Malignant	Benign	Total		Malignant	Benign	Total
FNAC+	0	0	0	**FNAC+**	4	5	9	**FNAC+**	52	9	61
FNAC-	11	101	112	**FNAC-**	0	0	0	**FNAC-**	0	0	0
Total	11	101	112	**Total**	4	5	9	**Total**	52	9	61

Table IV: Proportion of malignancy in Bethesda categories I, III and IV

Classification	Malignant (histology)	Benign (histology)	Total	Proportion of malignancy (%)
Bethesda I	1	1	2	50
Bethesda III	3	15	18	16.7
Bethesda IV	8	12	20	40
Total			40	

Figure 1 shows the proportion of each Bethesda category from the study. Bethesda I (inadequate) was yielded in only 2 FNACs.

In Table II, FNAC+/- indicates whether the FNAC detected malignancy (+) or not (-). FNAC (-) is equivalent to Bethesda II and FNAC (+) is equivalent to Bethesda V and VI. We found that the FNAC had a sensitivity of 83.6% [95% CI: 72.5%, 91.5%] and a specificity of 88.7% [95% CI: 81.4%, 93.8%]. FNAC performed well in comparison to histology in diagnosing true positive results and true negative results. The accuracy in this regard is 0.861 [95% CI: 0.808, 0.915] and it has a PPV of 81.2% [95% CI: 69.9%, 89.6%] and a NPV of 90.3% [95% CI: 83.2%, 95.0%] (Table II and Figure 2). These values were generated using Bethesda categories II, V and VI. Bethesda I is non-diagnostic and Bethesda III and IV are indeterminate, hence they were excluded when calculating these values.

Bethesda II reveals that FNAC has a specificity of 100%, an accuracy of 90% and a negative predictive value of 94.6%. These results reflect a high specificity and accuracy of FNAC in excluding thyroid cancer. Bethesda V reveals that FNAC has a sensitivity of 100%, an accuracy of 44%, and a PPV of 44.4%. The seemingly poor performance of FNAC in diagnosing thyroid nodules within the Bethesda V category is most likely due to the small sample size in this category. Bethesda VI reveals that FNAC has a sensitivity of 100%, an accuracy of 85% and a PPV of 85.2%. These results are consistent with FNAC being well-able to diagnose malignancy in thyroid nodules. This table indicates the raw statistics used to calculate overall diagnostic values as mentioned above. Each table displays the positive and negative results for both FNAC and histology. These tables ultimately generated true positive and negative values.

Bethesda I are thyroid nodules that were non-diagnostic or inadequate. According to the data received, only 2 FNAC results were classified into this category. The data translated to a 50% proportion of malignancy within this category due to one histology result being malignant and the other one being benign. Bethesda III are thyroid nodules that reveal atypia of undetermined significance or follicular lesion of undetermined significance. According to the data received, only 3 FNAC results out of the total of 15 results were diagnosed as a malignancy by histology. This translates to a 16.7% proportion of malignancy within this category. Bethesda IV are thyroid nodules that are a follicular neoplasm or suspicious for a follicular neoplasm. According to the data received, only 8 FNAC results out of the total of 20 results were diagnosed as a malignancy by histology. This translates to a 40% proportion of malignancy within this category (Table IV).

The risk of malignancy (ROM) as determined by the FNAC per Bethesda category in our study was compared to the ROM in Bethesda 2023 and reported in Table V.

Table V: Risk of malignancy in the present study and Bethesda 2023

Bethesda classification	ROM in our study (%)	ROM in Bethesda 2023 % (range)
Bethesda I (Non-diagnostic/unsatisfactory)	50	13 (5–20)
Bethesda II (benign)	0	4 (2–7)
Bethesda III (atypia of undetermined significance, or follicular lesion of undetermined significance)	16.7	22 (13–30)
Bethesda IV (follicular neoplasm or follicular carcinoma)	40	30 (23–34)
Bethesda V (suspicious for malignancy)	44.4	74 (67–83)
Bethesda VI (malignant)	85.2	97 (97–100)

Table VI: Nodule size and FNAC malignancy detection rate

Nodule Size	FNAC +	FNAC -	Total	FNAC malignancy detection rate (%)
< 4 cm	25	4	29	86.2
> 4 cm	30	6	36	83.3
Total	55	10	65	

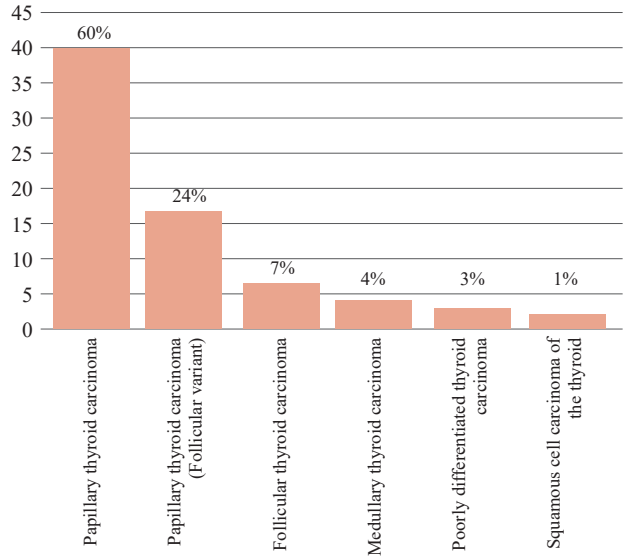

Figure 3: Prevalence of thyroid cancer subtype

FNAC malignancy detection rate was assessed for respective nodule sizes, with 4 cm being the cut-off size. In our study, smaller nodule sizes yielded slightly higher malignancy detection rates (Table VI).

FNAC detected malignancy at the highest rate for the rare thyroid cancers i.e medullary thyroid carcinoma, poorly differentiated thyroid carcinoma etc. It had a high malignancy detection rates i.e. > 80% for papillary thyroid carcinoma and the follicular variant of papillary thyroid carcinoma (Table VII).

Figure 3 shows the prevalence of each subtype of thyroid carcinoma, with papillary thyroid carcinoma being the most prevalent in this population.

Discussion

Our study described the sensitivity, specificity, and accuracy of FNAC when diagnosing thyroid cancer as being 83.6%, 88.7%, and 0.861, respectively. The PPV was shown to be 81.2%, the NPV was 90.3%, and the likelihood ratio generated was 7.39, which reflects a high diagnostic value. The FNAC performed much better in our study compared to the study done by Conradie et al., which found the diagnostic accuracy of FNAC to be 74% with a sensitivity of 73% and specificity of 74%.[3] Their PPV was 67% and NPV was 79%.[3] The possible reason for the discordance between these two similar studies is the fact that our study was performed in a single centre with the standardised FNAC technique. The study which was done by Conradie et al. evaluated national data from multiple centres with varying FNAC techniques.

When we compared the diagnostic performance of the FNAC in our study, it was concordant with other international studies, with marginal differences. These studies include a study performed by Michael et al. in 2016, which found the NPV of their FNAC to be 91% in their cohort.[4] We found a similar negative predictive value of 89.4% in our study. Similarly, Osseis et al. reported comparable sensitivity rates and PPV of 89.31% and 78% respectively. However, they also reported markedly lower specificity, lower NPV and accuracy rates of 48.44%, 68.89%, and 75.89% respectively.[5] These differences could be explained by different FNAC techniques and sample sizes. In their retrospective study, they included 344 patients while we included 222 patients.

It is worth mentioning that our study showed a very impressive diagnostic rate for the FNAC. Only two FNAC were found to be non-diagnostic, making the non-diagnostic rate 0.9%. This is an excellent performance compared to the 38% non-diagnostic rate which was reported by Conradie et al.[3]

The highlight of our study was the ability of our FNAC to rule out thyroid cancer with such high certainty. For benign disease, i.e., Bethesda II, there were no false positive cases. We found an impressive specificity of 100% and NPV of 94.6%. The ROM for Bethesda II was 0%, compared to the much higher ROM of 20.5% which was reported by Conradie et al. For Bethesda V, we found five out of nine cases, i.e., 55%, to be falsely positive. This means that these patients were subjected to unnecessary surgery. This highlights the need for molecular testing to further characterise the nodules in this category as recommended by the third edition of the Bethesda classification which was published in 2023.

When we compared the ROM for the different Bethesda categories in our study to the third edition of the Bethesda classification, we found similar rates for Bethesda III and IV. The ROM in our study for Bethesda category VI was lower, i.e., 85% compared to 97–100% in the 2023 Bethesda classification.

We found the prevalence of thyroid cancer at CMJAH to be 35.6%. This was similar to the local study done by Conradie et al., which found the prevalence of thyroid cancer to be 35.8%.[3] Similarly, Osseis et al. found the prevalence of thyroid cancer to be 38.67% and Michael et al. found it to be 40%.[5]

Table VII: FNAC malignancy detection rates for each histological subtype

Histological Subtype	FNAC +	FNAC -	Total	FNAC malignancy detection rate (%)
Papillary thyroid carcinoma	34	6	40	85
Papillary thyroid carcinoma (follicular variant)	14	2	16	87.5
Follicular thyroid carcinoma	2	3	5	40
Medullary thyroid carcinoma	3	0	3	100
Poorly differentiated thyroid carcinoma	2	0	2	100
Squamous cell carcinoma of the thyroid	1	0	1	100
Total	56	11	67	

The most common subtype in our study was papillary thyroid carcinoma, with the follicular variant being the most common. The malignancy detection rate by the FNAC was 85% and 87.5% respectively. Similarly, Conradie et al. found papillary thyroid carcinoma to be the most common subtype accounting for 61.8% of the histologically confirmed thyroid cancers when they evaluated the clinicopathological landscape of thyroid cancer in SA.[6]

We were able to highlight other factors influencing the diagnostic performance of FNAC, which include the size of the nodule used for the FNAC (whether < 4 cm or > 4 cm). We used the nodule size equal to 4 cm as the cut-off size as this indicates T3a in the American Joint Committee on Cancer. We found that the FNAC had slightly higher malignancy detection rates for nodule sizes < 4 cm, i.e., 86.2% compared to nodule sizes > 4 cm, i.e., 83.3%. This is corroborated by a study conducted in 2016 by Do Hoon Koo et al., which showed reduced diagnostic accuracy for thyroid nodules larger than 4 cm. As nodule size can deter accuracy of FNAC, it is an important component to highlight in our own study.[7]

In our study, the high proportion of thyroidectomies (94%) compared to lobectomies (6%) was noted. Our hospital (i.e., CMJAH) still performs mostly thyroidectomies for well-differentiated thyroid cancers of follicular cell origin, e.g., papillary thyroid cancer. The reasons for this include the fact that it is difficult to monitor for persistent or recurrent disease post-lobectomy in our setting and patients often get lost to follow-up. We perform diagnostic lobectomies for indeterminate nodules, i.e., Bethesda III and IV, and therapeutic lobectomies for Bethesda II with a clear indication for surgery, e.g., obstructive symptoms, cosmesis, etc., provided that the contralateral lobe is normal on neck ultrasound.

Further research is required to improve the diagnostic value of the indeterminate categories (i.e. Bethesda III and IV). Nikiforov et al. conducted a prospective study in 2009, whereby they aimed to test the feasibility and role of molecular testing of tumour-specific mutations in improving FNAC in diagnosing thyroid cancer. They found 97% of tumour mutation positive nodules to correlate with malignancy, as postsurgery.[8]

Limitations

This is a single centre study conducted retrospectively. Much of the additional limitations of the study are owing to human error in recording the original data at CMJAH. In some patients FNAC was performed and not recorded. A few patients were unable to be found on the NHLS online server. The sizes of the thyroid nodules were not reported or indicated in some patients' records. Selection bias is a possibility due to convenience sampling being used by the research team.

Conclusions

Our study findings indicate that FNAC has a high sensitivity, specificity and accuracy in diagnosing thyroid cancer preoperatively. When performed using a standardised technique, it results in far fewer non-diagnostic results, i.e., Bethesda category I. Therefore, it should continue to be used in the diagnostic workup of thyroid nodules at CMJAH. We recommend further research in this subject to be conducted in other institutions in SA, perhaps focusing on the indeterminate Bethesda categories, i.e., III, VI and V.

Acknowledgements

The authors would like to acknowledge the contributions of Drs Gabriella Do Pinheiro and Dr Clarence Yah in the initial student project.

Conflict of interest

The authors declare no conflict of interest.

Funding source

No funding was required.

Ethical approval

Ethical approval was received from the Human Research Ethics Committee of the University of the Witwatersrand (Ref: M221061). Further approval was received from the CEO of Charlotte Maxeke Johannesburg Academic Hospital.

ORCID

S Pillay https://orcid.org/0009-0000-7574-861X
T Ntrama https://orcid.org/0009-0001-5787-0328
G Muthen https://orcid.org/0009-0001-1008-5285
C Trainer https://orcid.org/0009-0008-6731-4414
Y Varachia https://orcid.org/0009-0001-2387-1621
M Mmutlane https://orcid.org/0009-0006-1086-0460
T Luvengho https://orcid.org/0000-0002-2901-1809
N Sishuba https://orcid.org/0000-0003-2333-8867

REFERENCES

1. Tsao GJ, Orloff LA. Clinician-performed thyroid ultrasound-guided fine-needle aspiration. Otolaryngol Clin North Am. 2014;47(4):509-18. https://doi.org/10.1016/j.otc.2014.04.005.
2. Aydoğan Bİ, Şahin M, Ceyhan K, et al. The influence of thyroid nodule size on the diagnostic efficacy and accuracy of ultrasound-guided fine-needle aspiration cytology. Diagn Cytopathol. 2019;47(7):682-7. https://doi.org/10.1002/dc.24170.
3. Conradie W, Baatjes K, Luvhengo T, et al. Performance of thyroid fine-needle aspiration biopsy in a low- and middle-income country. Acta Cytol. 2024;68(4):301-8. https://doi.org/10.1159/000539153.
4. Michael CW, Kameyama K, Kitagawa W, Azar N. Rapid on-site evaluation (ROSE) for fine needle aspiration of thyroid: Benefits, challenges and innovative solutions. Gland Surg. 2020;9(5):1708. https://doi.org/10.21037/gs-2019-catp-23.
5. Osseis M, Jammal G, Kazan D, Noun R. Comparison between fine needle aspiration cytology with histopathology in the diagnosis of thyroid nodules. J Pers Med. 2023;13(8):1197. https://doi.org/10.3390/jpm13081197.
6. Conradie W, Luvhengo T, Lübbe JA, et al. The clinicopathological landscape of thyroid cancer in South Africa: A multi-institutional review. World J Surg. 2024;48(12):2863-70. https://doi.org/10.1002/wjs.12353.
7. Koo DH, Song K, Kwon H, et al. Does tumour size influence the diagnostic accuracy of ultrasound-guided fine-needle aspiration cytology for thyroid nodules? Int J Endocrinol. 2016;2016:3803647. https://doi.org/10.1155/2016/3803647.
8. Nikiforov YE, Steward DL, Robinson-Smith TM, et al. Molecular testing for mutations in improving the fine-needle aspiration diagnosis of thyroid nodules. J Clin Endocrinol Metab. 2009;94(6):2092-8. https://doi.org/10.1210/jc.2009-0247.