The Accuracies of Fine-Needle Aspiration Biopsy without Ultrasound Guidance versus Frozen-Section Examination of Thyroid Nodule Diagnosis in Teaching Hospital Universitas Sumatera Utara: Single-Center Experience

Dedy Hermansyah, Denny Rifsal Siregar, Erwin Sahat Hamonangan Siregar, Iqbal Pahlevi Adeputra Nasution, Arjumardi Azrah Khalir Harahap, Naufal Nandita Firsty, Kamal Basri Siregar

1Division of Oncologic Surgery, Department of Surgery, Faculty of Medicine, Universitas Sumatera Utara, Medan, Indonesia; 2Department of Surgery, Faculty of Medicine, Universitas Sumatera Utara, Medan, Indonesia; 3Division of Pediatric Surgery, Department of Surgery, Faculty of Medicine, Universitas Sumatera Utara, Medan, Indonesia; 4Clinical Medical Student, Faculty of Medicine, Universitas Sumatera Utara, Medan, Indonesia

Abstract

BACKGROUND: The urgency to determine the most applicable methodological approach in thyroid nodules diagnosis workup is unquestionably important due to a significant increase in its prevalence globally. To date, fine-needle aspiration biopsy (FNAB) is the most recommended procedure to diagnose preoperatively even though the certain role of frozen-section examination (FSE) intraperatively were significant as well to be administered at some circumstances. 

AIM: This study aims to measure the diagnostic accuracies of FNAB and FSE and determine the role of each procedure in the current thyroid nodules' workup.

MATERIALS AND METHODS: This retrospective study was conducted at teaching hospital of Universitas Sumatera Utara, Indonesia. We included and investigated both FNAB and FSE in 106 patients; the results were compared to the final histopathological report, in which statistical analyses were conducted.

RESULTS: Most of the nodules are malignant after histopathological confirmation. From the fourfold contingency table, we were able to subtract several diagnostic parameters of the procedures. The FSE substantially disclosed better performance in sensitivity, specificity, positive predictive value (PPV), negative predictive value, and remarkably the accuracy. After several in-depth analyses, it is apparent that a higher rate of false-negative in FNAB is inducing its substandard performance, especially when the guidance of ultrasonography was not applied in this study. Furthermore, an analysis toward the concordance between both procedures results is revealing a significant improvement on almost every parameter, indicated by 100.0% specificity and PPV.

CONCLUSION: Although the outcomes in this study are not corresponding to the current guidelines, we acknowledged several limitations in this study. However, the role of FSE in thyroid nodule workup should be reconsidered since we found that the procedure is exhibiting a potential to assist the current method, even though the influence of its certain pitfalls remained questionable.

Introduction

The prevalence of thyroid nodules globally is reported to be progressively more common particularly due to the advances in the available approaches; and although most of the nodules are generally benign, the risk of its being malignant should always be perceived in the diagnostic workup. The recent epidemiological investigation of thyroid nodules is relatively challenging to be reported and to be identified on some databases as well. It is suspected that 4–7% of adult populations (mostly women) are presenting at least one palpable thyroid nodule, in which the malignant lesion may be diagnosed among approximately 5% of those populations [1], [2], [3]. Kou et al., in its report of thyroid nodule prevalence among 9146 adults in China, revealed that the rate is gradually higher with age, even reaching 72.4% in populations aged 70–89 years old along with the confirmation of its higher number in females [4]. Regardless from its diagnosis, the clinical presentation of each patient is diverse and may resemble each other, ranging from asymptomatic and appear to be benign as its small and solitary, to symptomatic and large nodules. Hence, it is unquestionably important to identify and distinguish the malignancy to determine the next possible step in the nodules’ management, whether it is still appropriate to directly order definite approach, that is, surgical treatment or the nodules are located in uncertainty range – hence, further utility of diagnostic test should be considered.

To date, fine-needle aspiration biopsy (FNAB) or cytology (FNAC) with or without ultrasound (US)-guidance and frozen-section examination (FSE) are
the leading procedures for histopathological analysis of thyroid nodule, even though current guidelines by AACE/AME task force is highly recommending US guidance for FNAB, primarily toward the high US risk thyroid lesions ≥10 mm, or intermediate and low risk with higher nodule size (≥20 mm, with supporting history for low-risk US results). Followingly, the FSE may be considered if the FNA results are suspiciously malignant or exhibited with inadequate cellarularity; hence, further workup is necessary; and at that particular part, the FSE may have a role to confirm the diagnosis intraoperatively [5]. The guidelines by European Thyroid Association also had suggested the routine use of FNAB as it plays a central role in the workup process, even though the execution should be selective as the risk of the results being incorrect is apparent [6]. Furthermore, the American Thyroid Association consensus in 2015 also mentioned that FSE disclosed the most helpful utility in the classic papillary thyroid carcinoma, but not in the special settings, for example, indeterminate lesions thus credited as its main reason to be abandoned in several guidelines, as proved by low number of recommendation to undergo the procedure [7]. In this study, we aim to elaborate the utilization of FNAB and FSE in clinical settings of thyroid nodules diagnostic workup, especially in terms of whether the US guidance in FNAB is in fact should be routinely applied, which, hence, justify the decision to abandon unguided procedures in most centers globally or even the FSE may be abandoned due to its unfavorable recommendation in several international guidelines.

We acknowledge the fact that US guidance in FNAB procedure of thyroid is fundamental in current guidelines; therefore, this study will also attempt to prove how pivotal the US guidance in general as exhibited in our center investigation by comparing the unguided FNAB workup to the considerably abandoned procedures in routine practice, that is, FSE. Indirectly, our study may also estimate the current impact of FNAB-without US guidance in insufficient of tools settings, as the availability of US in every center is eventually challenging to be thoroughly confirmed in most developing countries, for example, Indonesia. Moreover, it is also our objective to provide an eligible data for oncologic sciences possible review, since the issues of under treatment as caused by underdiagnosis or vice versa are definitely burdensome.

**Materials and Methods**

This retrospective study was conducted from the medical record of the teaching hospital of Universitas Sumatera Utara in Medan, Indonesia from 2018 to 2019 with confirmed benign or malignant thyroid nodules and treated operatively (lobectomy, isthmulobectomy, and total thyroidectomy). The FNAB procedure was undertaken preoperatively without USG guidance, followed by standard FSE procedures intraoperatively, and histopathological analysis postoperatively for each participant in surgical oncology division. We included a total of 106 individuals with complete medical record of the administered intervention as admitted to the department of surgical oncology in the same institutions. The specimens were submitted for diagnostic confirmation to the Department of Surgical Pathology, Universitas Sumatera Utara General Hospital.

The FNAB examination results were categorized into two major classifications according to its final cytology outlines, which are: benign cytology (Thy I-IV), consisting several conditions, for example, nodular goiter, thyroiditis, and the other non-malignant diagnosis, in which ‘generally’ labeled as “benign smear. Conversely, the ‘malignant smear’ is applied to annotate confirmed malignant cytology (Thy V-VI) results which include the papillary, undifferentiated (anaplastic), follicular, medullary, or mixed thyroid carcinoma. The latter statements of classifications were based on Bethesda system for thyroid cytopathology reporting (Anand et al., 2020). The similar classification method was also implemented in FSE procedures, which principally assigned into both benign and malignant lesion, representing non-malignant findings, and carcinoma presence, respectively. Accordingly, the initial report of both detection method was compared to the final histopathological interpretation, which certainly posed as the gold standard in the diagnosis of nodular thyroid diagnosis workup; corresponding to the previous malignancy-confirmation methods hence determined as either benign or malignant results. All participants were registered to underwent all diagnostic examination; therefore, the comparison can be adjusted thoroughly with bias minimalization as the result of different procedure.

Therefore, the calculation of both cytologic FNAB and FSE in term of diagnostic accuracies was assessed by the comparison toward the histopathological report at the end of each thyroid nodule evaluations’ session. The data were represented by the following parameters, for example, (1) sensitivity (Sen); (2) specificity (Sp); (3) positive predictive value (PPV); (4) negative predictive value (NPV) and accuracy (Acc). The statistical analysis was performed using IBM SPSS Statistics version 24.0. The analytical insight and method of this study (to deliver and delineate) the results of this study were heavily influenced by preceding investigations worldwide, even though we included several innovations in data presentations [8], [9].

https://oamjms.eu/index.php/mjms/index
Results

The participants’ demographics are demonstrated in Table 1, encompassing several crucial information, for example, gender, mean age (in years; as documented in its first presentation), and the final diagnosis as outlined by the histopathological report.

Table 1: Characteristics of the final report from each procedure.

| Table 1: Characteristics of the final report from each procedure. Note that every participant in this study underwent all diagnostic procedures |
|--------------------------------------------------|
| Diagnostic conclusion | Number of patients | Percentage of total patients | Mean age (years) | Gender (female: male) |
|------------------------|--------------------|-----------------------------|-----------------|----------------------|
| Total                  | 55                 | 23                          | 55              | 12                   |
| FNAB report            | 55                 | 23                          | 55              | 12                   |
| Benign nodule          | 14                 | 13.2                        | 44.9 ± 12.2     | 50:1                 |
| Malignant nodule       | 92                 | 86.8                        | 51.0 ± 13.0     | 27:1                 |
| Benign cytology        | 47                 | 44.3                        | 48.2 ± 12.9     | 68.1                 |
| Malignant cytology     | 59                 | 55.7                        | 52.1 ± 12.9     | 17:1                 |
| FSE report             | 21                 | 19.8                        | 47.2 ± 11.9     | 60:1                 |
| Benign FSE             | 85                 | 80.2                        | 51.2 ± 13.1     | 17:1                 |

FNAB: Fine-needle aspiration biopsy, FSE: Frozen-section examination.

As mentioned earlier, all 106 participants’ final reports from each procedure were compared to determine the diagnostic Acc. From the Tables 2 and 3, the detail of both procedure comparison along with its final statistical analysis can be observed; considering there are several remarkable dissimilarities and variability in the procedures’ diagnostic accuracies. Even though the 2 × 2 contingency table (first-half of Table 3) are representing the result of all undertaken examination method in this study, Table 2 provided better perspective to understand the measure of each procedure. The design of Table 2 was aimed to represent the rate of final diagnosis (basically a benign/malignant guidance) in both procedures, as well as depicting the diversion of our final reports. Accordingly, the measurement of each or even both diagnostic accuracies can be statistically clear, either the capability of single-procedure approach or the consequence of straight benign (or malignant) result in FNAB and FSE.

To exemplify the results, the FSE exhibited substantially better outcomes in all parameters as represented by higher diagnostic capacity especially in the Sn (91.3% vs. 62.0%) and NPV (61.9% vs. 25.5%) value, in which influenced by the higher rate of false-negative (FN) in FNAB procedures. Correspondingly, higher false-positive (FP) rate (although the number was not too significant in difference; 2 vs. 1 in FNAB and FSE, respectively) of the latter procedure also affect the Sp (92.9% vs. 85.7%) and PPV (98.8% vs. 96.6%) parameter, both displaying the FSE favorability. The dissimilarity in diagnostic performance also revealed by the Acc value in both procedure, which are 65.1% and 91.5% for FNAB and FSE, respectively.

Therefore, our study found that the FSE procedures had remarkable capability to determine an individual thyroid nodule test as a malignancy; and as the consequence, the frozen-section also disclosed better Acc in term of settling a true-positive result. Although the Sp or PPV superiority of FSE were not as apparent as the other two parameters, it is still noteworthy to mention that both procedures were able to identify the true-negative result, or distinguishing a FP result, hence may participate in the overall reduction of the over-treatment or exaggerated diagnostic workup which unquestionably should be avoidable. Conversely, as there was such significant difference in Sn and NPV results, the FSE procedures were expected to differentiate the FN results more favorably; considerably remarkable to avoid undertreatment case corresponding to the undetected malignancies issue in clinical oncologic workup settings.

Table 2: Rate of benign/malignant results in FNAB and FSE as compared to the final histopathological report.

| Table 2: Rate of benign/malignant results in FNAB and FSE as compared to the final histopathological report |
|--------------------------------------------------|
| Final diagnosis of evaluated procedures (FNAB/FSE) | Final histopathological report (n) | Total |
|--------------------------------------------------|----------------------------------|-------|
| Benign                                           | Malignant                        |       |
| FNAB                                             | 11                               | 12    | 23 |
| FSE                                              | 1                                | 23    | 24 |
| Total                                            | 14                               | 92    | 106 |

FNAB: Fine-needle aspiration biopsy, FSE: Frozen-section examination.

As documented in its first presentation), and the final histopathological report (basically a benign/malignant guidance) in both procedures, as well as depicting the diversion of our final reports. Accordingly, the measurement of each or even both diagnostic accuracies can be statistically clear, either the capability of single-procedure approach or the consequence of straight benign (or malignant) result in FNAB and FSE.

Further, concordance of these diagnostic accuracies analysis especially to only both benign and both malignant results in FNAB and FSE procedures (B/B and M/M) as mentioned in Table 2 will yield much favorable results; as exhibited by Sn, Sp, PPV, and NPV value of 82.1%, 100.0%, 100.0%, and 47.8%, respectively, in which significantly higher than single-diagnostic tool utilization. Even though we acknowledged the fact that this analysis only represents 73.6% of the total participants, the concordance of final diagnosis in FNAB and FSE is capable to disclose better diagnostic performance.

Discussion

The current diagnostic workup for thyroid nodules findings appeared to be vast in term of options, especially since the introduction of cytologic-based...
investigation, for example, FNAB or FNAC in the early 1930s by Martin and Ellis which provided better establishment for oncologic detection or malignancy determination (due to its procedure Sp to the most suspected part of the nodules by needle-insertion technique) [10], [11]. Recently, the utilization of ultrasonography (US) in FNAB has become the preference in thyroid nodules investigation as it allows an anatomical guidance to either palpable or non-palpable lesion. However, as mentioned in earlier section, the FNAB procedures in this study were not guided by US, even though we perceive that the role of US in this setting had conduct such profound change in thyroid diagnostics [12]. Therefore, we would like to address that the results of this study were eventually affected by the pre-acknowledged “modus operandi” to not apply US guidance FNAB. Conversely, the FSE is basically an intraoperative procedure, which enable rapid pathological report hence supporting the clinicians to diagnose the tissue samples (or thyroid nodules’ sample in this study) immediately based on its procedural findings, although several indeterminate variant of a nodule are easily misdiagnosed [13], [14].

According to several guidelines and task-force discussion outcomes, it is widely perceived that FNAB is the most accurate method for evaluating a suspected thyroid nodule-hence functioning as justifying tool to determine the necessity for further definitive management (along with the consideration to the patients’ clinical or laboratory analysis). However, the latter statement is generally acceptable if only the diagnostic procedures implemented the US guidance especially for some complex nodules or posteriorly located nodule; unquestionably increasing its Acc either anatomically or histopathological investigation in advance [12], [15], [16]. Our findings of FNAB (without US-guided) accuracies parameter were similarly reported in a recent study by Youssef et al., in 2020, in which stating the Sn, Sp, PPV, and NPV value of US-guided FNAB for 37.5%, 100.0%, 100.0%, and 77.27%, respectively; for a review purpose, the same study applied the utilization of Bethesda scoring method for its nodule status determination (I–IV for benign and V–VI for malignant) with p-value of 0.001 in the its diagnostic performance analysis [17]. Another study by Bozbıyık et al., in 2011, demonstrate similar value of accuracies parameter value to determine the diagnosis in large thyroid nodules as compared to Youssef et al. and our study. Bozbıyık et al. also had conducted a mini-systematic-review of a series which comparing a reliability of FNAB in diagnosing large thyroid nodules, with eventually similar findings with our results [18]. Yoon et al. also had revealed substantially higher diagnostic value of US-guided FNAB with 96.7%, 85.9%, 76.6%, 98.2%, and 89.4% for Sn, Sp, PPV, NPV, and Acc, respectively, supporting the premise of FNAB’s accuracies variability regarding its performance, regardless the role of US guidance in the procedures [19].

The comparison between FNAB and FSE also somehow lead to such debatable aspects of oncologic diagnostic forum worldwide. We believe that our study had provided an additional perspective to view this matter, even though some limitation should be acknowledged earlier in this section. The reports of how precise FSE in predicting a malignancy in a thyroid nodule sample are also strikingly diverse generally, but it is remarkable as well to mention that several studies, in which evaluating both FNAB and FSE disclosed similar result as we found-with FSE results are exhibiting better performance especially for Sn parameters. As we found, there was a substantial difference of both procedure in Sn analysis (91.3% vs. 62.0%, for FSE and FNAB, respectively). Guevara et al. also reported some comparable data of the Sn value, contracting 75.4% for FSE versus 31.9% for FNAB, although the same study mentioned both procedures are performing excellent (100%) in Sp and PPV parameters [20]. Similar report also had been provided by Kahmke et al., revealing the accuracies parameters of both procedures (FSE vs. FNAB) which are Sn (76.9% vs. 53.8%), Sp (67.9% vs. 74.1%), PPV (27.8% vs. 25.0%), and NPV (94.8% vs. 90.9%) [13]. Furthermore, an earlier report by Cetin et al. (2004) [21] and Huber et al., (2007) [22] had provided corresponding and relatively equivalent results, partially delineating the “inconsistent” results were consistently outlined internationally thorough decades; even though an “incomplete” final interpretation regarding these diagnostic parameters review is the FNAB procedure that had a tendency to performed substantially in Sn or NPV value, possibly influenced by its relatively higher FN results [21], [22].

The concordant results between FNAB and FSE may also prove beneficial to predict a diagnosis more accurately. Our study clearly stated if the latter outcomes were observed in a thyroid nodule sample, all diagnostic parameters particularly the value in which involving FP in its equation will perform significantly better. As a results, in this investigation the 100% Sp and PPV can be interpreted, only if the results of FNAB and FSE are in a concordance, or matching, for example, both benign or malignant-hence the final histopathological results will more likely to manifest the previously accordant diagnosis. Young et al., in 2011, had outlined this issue and confirmed that the concordant findings of FNAB and FSE were exhibited better accuracies (with 100% Sp and PPV as well); followed by higher Acc rate in which is not observed in this study [9]. Unfortunately, it is remarkably challenging to find the study which evaluated the concordance of both diagnostic procedures. Nevertheless, an earlier study from 2002 by Chang et al. reported that the Sn, Sp, PPV, NPV, and Acc of a concordant results are 92.3%, 100.0%, 100.0%, 97.4%, and 98.0%, respectively – the latter value was re-analyzed by us since the study only provided the outcomes of each procedure [23].

It should be acknowledged that each procedure is possessing different limitations, considering the
method of investigation is almost entirely different. Although FNAB as a procedure is frequently favorable since its reckoned as a safe and effective approach, several pitfalls also recognizable, for example, operators’ skill-based, variable specimen adequacy, and the procedure itself may alter the appearance of the nodules, as reported by Moon et al., in which elaborated that post-FNAB follicular adenoma nodule may be mistaken for medullary thyroid carcinoma since the pathologist observed an extensive fibrosis caused by the aspiration procedure [24], [25]. The higher FN results also had proved in this study as we “only” found 25.5% NPV of FNAB procedures. The existence of “gray zone” in follicular pattern lesion produced a difficult condition to rule out malignancy on cytology procedures; therefore, histopathological analysis should always be considered in latter circumstances. Zhu et al., in its systematic review, mentioned that specimen problems are leading the most common etiology for FN results in FNAB, while the interpretation error may lead to FP conclusion; notice how these two factors should be noted since it also had occurred in this study as well—especially when the US guidance was absent and it is somehow susceptible for some false-results [26], [27].

Our perspective toward the issue regarding the FSE routinely use is how this procedure progressively more uncommon to be implemented, since several factors, for example, sufficient experience from the pathologist (it is also necessary in FNAB) and technical processing factors (sample-related details) are generally important. Sanabria et al. had widely explained why it is appropriate to abandon FSE for routine use in thyroid gland neoplasm investigation, specifically in the follicular-based lesion. The review itself was focusing on low diagnostic performance and high rate of deferred results since the latter sample will falsely increase the Acc due to the preceding information or diagnostic review biases; and it is eventually hard to be avoided since it depends on several factors elaborated earlier (sample factors or operator-dependent) [14], [20], [28].

Therefore, according to our findings and after several consideration in the previous discussion, the comparison of both procedures is delineating the fact that FSE may still possess several core or even adjunctive capacity to assist the current recommendation of thyroid nodule diagnosis, either by US only or even FNAB as delineated in this study. Nevertheless, if the elaboration was solely constructed based on our findings, the role of US guidance in FNAB procedure is certainly massive enough to influence the outcomes in this study since the diagnostic performance of unguided-FNAB procedure is comparable or even relatively inferior to FSE, hence justifying the overall decision in major healthcare stakeholders to implement US guidance in routine FNAB to amplifying the accuracies.

Hence, as outlined in the international guidelines of thyroid nodule workup, we advised that the application of FNAB procedure should be accompanied by US as the absence of its guidance in diagnostic settings is highly influential. However, in the special settings, for example, unavailability of the guidance as we had experienced or at least simulated in this investigation since we aimed to prove the impact of unguided FNAB, the FSE procedures are plausibly reliable to be applied if the suspicious lesions were encountered intraoperatively. To the latter statements, we believe that the approach of thyroid nodule should be flexible but reasonable in practice, with some adjustment to the encountered settings are highly recommended, even though the role of international guidelines should account the decision at the first place. Furthermore, we believe that more studies are necessary to confirm those statements, as the analysis toward several outcomes, for example, concordant result between guided-FNAB and FSE, or even the comparison between guided- and unguided-FNAB as reviewed systematically and quantitatively should be conducted to enhance the evidence of FNAB utilization.

Conclusion

We acknowledged that the results are slightly unexpected from the recent recommendation for thyroid nodule workup and guidelines yet technically explainable. Although, the inconsistencies of both procedure’s report were unquestionably influencing the background and reason to conduct this study at the first place, in which we tried to resolve. We, still, encourage the further investigation of this particular issue considering the FSE itself is performing significantly better than unguided-FNAB hence still supporting the utilization of FSE in daily practice to assist the workup procedures, although it should be realized that US guidance is standing integral in the FNAB procedures preoperatively. Especially, the findings of concordance results between unguided-FNAB and FSE results will produce a better diagnostic performance; even though we realized the cost- or resources-effectiveness in that recommendation is perhaps mediocre to be fair, and following recommendation to delineate the most effective approach should always be considered.

References

1. Hong A, Valderrama E, O’Reilly B, Lanzkowsky P, Friedman D, Gandhi M, et al. A thyroid nodule. Child Hosp Q. 1999;9:33-5.
2. Jiang H, Tian Y, Yan W, Kong Y, Wang H, Wang A, et al. The prevalence of thyroid nodules and an analysis of related lifestyle factors in Beijing communities. Int J Environ Res Public Health. 2018;15(4):1-11. https://doi.org/10.3390/ijerph15040442 PMid:27110805
3. Parsa AA, Gharib H. Epidemiology of thyroid nodules. In:
