

Determination of the ideal sampling technique to reduce repeated procedures: a comparative study including 393 fine-needle aspirations for thyroid nodules

Fatih UZUNKAYA^{1*}, Ahmet ÖZDEN²

¹Department of Radiology, Faculty of Medicine, Ondokuz Mayıs University, Samsun, Turkey

²Section of Radiology, Dr. Münif İslamoğlu State Hospital, Kastamonu, Turkey

Received: 10.01.2016 • Accepted/Published Online: 27.04.2016 • Final Version: 27.02.2017

Background/aim: Fine-needle aspiration biopsy is an established method for the evaluation of thyroid nodules, but it has not been standardized worldwide yet. Adequacy of the aspirations is affected by several factors. The aim of this study is to determine the main factors affecting the adequacy and to suggest a procedural technique expected to reduce repeated procedures.

Materials and methods: A total of 393 aspiration procedures performed using either 22-gauge or 27-gauge needles were included in the study. The samplings were classified as inadequate or adequate according to the cytopathological reports, and results were compared.

Results: The rate of adequate samplings was higher in the 27-gauge group and the difference was statistically significant. Neither the size of nodules nor the number of slides used for smearing affected the adequacy. There was not a statistically significant relation between the needle size and the nodule size or the number of slides in terms of adequacy.

Conclusion: Needle size is an important factor that affects the adequacy of samplings. The nodule size and the number of slides do not affect the adequacy. However, bloody and thicker smears are difficult for pathologists to evaluate and result in inadequacy.

Key words: Thyroid nodule, fine needle aspiration, adequacy

1. Introduction

In the adult population thyroid nodules are very common and they are detected in more than half of all ultrasound (US) examinations (1). Thyroid malignancy is relatively rare and most thyroid nodules are benign (1,2). Fine-needle aspiration (FNA) biopsy is an established method for the evaluation of thyroid nodules so as to differentiate benign from malignant (3–9). By using FNA, patients with benign nodules are appropriately triaged and the rate of unnecessary thyroid surgery is reduced (10). Adequacy of sampling is crucial for cytopathological evaluation for FNA (10). In the literature, there is a large range in adequacy rates, from 0.4% to 40.7%, and approximately 20% of the procedures are inadequate/nondiagnostic (4,11). An inadequate thyroid FNA is an annoying issue both for patients and physicians. There are some factors that might affect the rate of adequacy of samplings (2,4,5,9,11–14). There are different procedural techniques of operators, essentially dependent on resources available and habits. Therefore, FNA techniques for the thyroid have not yet been standardized (4,9). In this study, we aimed to determine the main factors that affect the rate of adequacy of samplings and to suggest a procedural technique

expected to decrease repeated procedures in the absence of immediate on-site cytopathological analysis.

2. Materials and methods

The medical records of US-guided FNA procedures for thyroid nodules performed in our radiology unit between October 2013 and April 2015 were retrospectively evaluated. The study included 376 patients (78 males, 298 females; aged 15–86 years; mean age: 50.9 years). A total of 393 aspiration procedures without an on-site cytopathological analysis were performed in this period. Patients were referred to us for FNA by their primary care physicians. No patient had a history of thyroid surgery.

Procedures were performed by two radiologists using different techniques regarding needle size, number of passes while sampling, and number of glass slides used for smearing the samples. All the cytopathological reports of aspiration procedures were analyzed. The reports were classified as either inadequate or adequate for diagnosis; the latter consisted of benign findings, atypia of undetermined significance or follicular lesion of undetermined significance, follicular neoplasm or suspicious for a follicular neoplasm, suspicious for malignancy, and

* Correspondence: drfatihuzunkaya@gmail.com

malignant according to the categories established by the Bethesda System for Reporting Thyroid Cytopathology (10). According to this system, every thyroid FNA was first evaluated for adequacy by the two pathologists before a specific diagnosis was defined. Inadequate samples were reported as nondiagnostic. This category was applied to samples that were unsatisfactory/inadequate/insufficient owing to obscuring blood, overly thick smears, or an inadequate number of follicular cells. For a thyroid FNA sample to be satisfactory/adequate for evaluation, at least 6 groups of benign follicular cells were required, each group composed of at least 10 cells. There were several exceptions to the numeric requirement of benign follicular cells. Any sample that contained abundant colloid was considered adequate, even if 6 groups of follicular cells were not identified. Whenever a specific diagnosis (e.g., lymphocytic thyroiditis) could be rendered and whenever there was any atypia, the sample was adequate for evaluation. Samples that consisted only of cyst contents (macrophages) were included in the nondiagnostic category. We set two major groups for the reports, including nondiagnostic and the remaining diagnostic group. After this grouping, we assigned samplings for both radiologists as inadequate for cytological material corresponding to nondiagnostic results, and adequate for cytological material corresponding to diagnostic results. We then evaluated the data in order to compare each sampling technique presented by the two radiologists for the purposes of the study.

The study was approved by the local ethics committee. Written informed consent was obtained from all participants.

2.1. Sampling technique

Both radiologists informed all patients about the aspiration procedure before the examination. The patients were questioned about hemorrhagic diathesis and anticoagulation therapy. Patients with hemorrhage risk were not approved.

All patients were randomly examined with three different US devices (Mindray, DC-7T, Shenzhen, China; Esaote, MyLab 70 XVision, Genoa, Italy; and Hitachi Aloka, F37, Tokyo, Japan). To localize the nodule to be sampled, a preliminary examination was performed by 7.5–10 MHz multifrequency linear-array probe. If a patient had multiple nodules, the most suspicious one was chosen according to malignancy features such as marked hypoechogenicity relative to surrounding strap muscles, microcalcifications, local invasion, and an ill-defined and irregular margin (1). In nodules without specific malignancy features, the largest nodule in size was targeted for aspiration.

After evaluation, the puncture side of the neck was cleansed with 10% povidone-iodine solution. A freehand biopsy technique without local anesthesia was used in

all aspiration procedures. Nodules were targeted in the axial plane. In the setting of a close neighborhood of large vessels or the trachea, a convenient oblique plane and short axis approach was chosen. Punctures were generally performed from the medial side of the probe in order not to pass the sternocleidomastoid muscle.

There was some technical diversity between the two radiologists performing aspiration procedures. Radiologist I, representing the first sampling technique, used 22-G needles attached to 5 or 10 mL syringes in samplings. With this technique, the aspiration procedure was performed by moving the needle back and forth within the nodule and approximately four or five oscillations with an equivalent vacuum pressure of the syringes via only one puncture. Samples acquired by this technique were smeared using a nonpredetermined number of slides suitable to the amount of samples.

Radiologist II, representing the second sampling technique, used 27-G needles attached to 2.5 mL syringes in samplings. With this technique, aspiration procedures were performed in almost the same way. However, there was a difference in determining the number of punctures. While Radiologist I performed only one puncture each time, Radiologist II arranged the number of punctures according to a targeted number of slides. The target was to obtain five to ten slides with rarefied smears. Hence, the number of punctures ranged from one to three, mostly being two. In some cases, the targeted number of slides could not be achieved because of an excess or paucity of samples, and the obtained slides were sent for cytopathological evaluation.

All of the smears acquired with both techniques were air-dried. The majority of the sampled nodules were solid or predominantly solid, and a few nodules were predominantly cystic. In the case of cystic nodules, the solid component was sampled. The smears were examined by two pathologists blinded to the study after the samples arrived to pathology laboratory, not on-site.

2.2. Statistical analysis

Data were analyzed with commercially available software (SPSS; SPSS Inc., Chicago, IL, USA). Continuous variables are presented as means, and categorical variables are presented as frequency and percentage. The chi-square test was used for statistical analysis, and $P < 0.05$ was considered to indicate a significant difference.

3. Results

In all cases, no major complication was encountered during or after the procedures. Hemorrhage into a cystic nodule occurred in one procedure performed using a 27-G needle and it limited itself after compression to the puncture side of the neck.

The number of aspiration procedures in the cases was as follows: one in 361 patients, two in 13 patients, and three in two patients (total of 393 aspiration procedures). The aspiration procedure was repeated in 11 cases due to nondiagnostic results.

Aspiration procedures of 215 (54.7%) cases were performed using a 22-G needle, and 178 (45.3%) were performed using a 27-G needle. The distribution of cytopathological results with respect to the samplings performed using a 22-G needle is as follows: inadequate samples for diagnosis in 76, benign in 124, atypia of undetermined significance or follicular lesion of undetermined significance in 12, suspicious for malignancy in one, and malignant in two. The distribution with 27-G needles is as follows: inadequate samples for diagnosis in 37, benign in 121, atypia of undetermined significance or follicular lesion of undetermined significance in 18, and suspicious for malignancy in two.

The rate of adequacy of samplings according to needle size is presented in Table 1. There was a statistically significant difference between the needle sizes in terms of sampling adequacy, in favor of the 27-G group ($P = 0.01$).

Size of the nodules ranged from 4 to 60 mm (mean: 20.27 mm). The rate of adequacy of samplings according to nodule size is presented in Table 1. There was not a statistically significant difference regarding nodule size in terms of sampling adequacy ($P = 0.794$).

Number of slides used for smearing samples ranged from 2 to 16. For the first operator (using 22-G needles) the number of slides was minimum 4, maximum 16 (mean: 8.72). For the second operator (using 27-G needles) it was minimum 2, maximum 12 (mean: 7.37). The rate of adequacy of samplings according to number of slides is presented in Table 1. There was not a statistically

significant difference regarding number of slides in terms of sampling adequacy ($P = 0.098$).

Distribution of adequate samplings according to needle and nodule size was evaluated individually. This distribution is presented in Table 2. There was not a statistically significant relation between the needle and nodule size in terms of sampling adequacy ($P = 0.648$).

Distribution of adequate samplings according to needle size and number of slides was also examined individually. This distribution is presented in Table 3. There was not a statistically significant relation between the needle size and number of slides in terms of sampling adequacy ($P = 0.094$).

4. Discussion

An inadequate FNA procedure is an annoying issue. A nondiagnostic report causes frustration for the referring clinician, the radiologist, the pathologist, and for the patient above all. It means added costs to the health care system for repeat testing and extra loads on departments such as radiology and pathology. Repeating aspirations also raises the patient's anxiety. In addition, nondiagnostic FNA results should not be regarded as simply benign because 2%–14% of nondiagnostic results are eventually proved to be malignant by surgery (11). To increase the rate of adequate samplings in order to avoid repeated aspiration is the ultimate goal of operators; nevertheless, there is a limit. The rate of adequacy varies widely in the literature (2,4,5,9,10,11,13,15). The range published by Ogori et al. (15) is between 76% and 95%. An approximate rate could be accepted as 80% (4).

There are some factors that might affect the rate of adequacy of samplings (2,4,5,9,11–14). It may be affected by the level of operator experience, accuracy of

Table 1. The rate of adequacy of samplings according to needle size, nodule size, and number of slides.

| | Number of samplings (%) | Inadequate (%) | Adequate (%) | P-value |
|------------------|-------------------------|----------------|--------------|---------|
| Needle size (G) | | | | |
| 22 | 215 (54.7) | 76 (35.3) | 139 (64.7) | |
| 27 | 178 (45.3) | 37 (20.8) | 141 (79.2) | 0.001 |
| Nodule size (mm) | | | | |
| <10 | 9 (2.3) | 2 (22.2) | 7 (77.8) | |
| 10–30 | 334 (85) | 95 (28.4) | 239 (71.6) | |
| >30 | 50 (12.7) | 16 (32) | 34 (68) | 0.794 |
| Slides (number) | | | | |
| <5 | 27 (6.9) | 12 (44.4) | 15 (55.6) | |
| 5–10 | 306 (77.9) | 81 (26.5) | 225 (73.5) | |
| >10 | 60 (15.3) | 20 (33.3) | 40 (66.7) | 0.098 |

Table 2. The distribution of adequate samplings according to needle and nodule size.

| | Number of samplings (%) | | P-value |
|------------------|-------------------------|-----------|---------|
| | 22-G | 27-G | |
| Nodule size (mm) | | | |
| 10 | 8 (44.4) | 10 (55.6) | |
| >10 | 131 (50) | 131 (50) | 0.648 |

localization of the lesion and the needle tip, method of guidance (US or palpation), intrinsic characteristics of the nodule, number of needle punctures, needle size, sampling technique (capillary action vs. aspiration), capability for immediate on-site cytopathological analysis, and many other factors (13). Recently, Lee et al. (11) published an article summarizing factors influencing cytological adequacy, based on a literature review. Many other studies have been done on this issue; nevertheless, there is not a well-accepted guideline for an ideal procedural technique for FNA.

We think it would be valuable to develop a consensus statement for the procedure, despite the fact that the procedure is essentially dependent on available resources and habits. In this study, our aim was not to design a multifactorial analysis; rather, we aimed to evaluate the contribution of the most easily modifiable and the most objectively definable variables in our daily practice to the adequacy of sampling. Thus, we expected to find solutions to decrease our frequency of repeated aspiration based on available resources and to contribute to the establishment of an ideal procedural technique. We decided that the most easily modifiable variable in our setting was needle size. We also determined the most objectively definable variable in our setting to be nodule size. We evaluated whether the number of slides used for smearing the samples is a factor affecting sample adequacy or not, for the first time as far as we know.

There is currently no universally accepted guideline for needle size in FNA techniques (5). Current FNA needles range from 21 to 27 G (4). Degirmenci et al. (2) reported that selecting smaller-sized needles decreased the rate of inadequate material in cytological examination. In a previous study, it was reported that: "With more cells aspirated, the chance of sufficiency for diagnosis increases. Large-bore needles lead to more cellular material being aspirated but more bloody specimens that may interfere with cytopathological interpretation. Small-bore needles may result in too few cells for diagnosis" (5). We have confirmed some of this knowledge with our study. Our results with 27-G needles was concordant (79.2%) with

Table 3. The distribution of adequate samplings according to needle size and number of slides.

| | Number of samplings (%) | | P-value |
|-----------------|-------------------------|------------|---------|
| | 22-G | 27-G | |
| Slides (number) | | | |
| <5 | 6 (40) | 9 (60) | |
| 5–10 | 107 (47.6) | 118 (52.4) | |
| >10 | 26 (65) | 14 (35) | 0.094 |

the literature in terms of adequacy, whereas the adequacy rate with 22-G needles was somewhat lower (64.7%). The difference between needle sizes was also statistically significant ($P = 0.01$). Moreover, the size of the needle might also affect the comfort of patient (4,5,16). In terms of interventional procedures, an adequate and comfortable sampling technique with less painful instruments is the state of the art. Hence, we can easily declare that the 27-G needle for aspiration is an efficient and comfortable tool of choice.

In our study, the lowest adequate sampling rate was in nodules larger than 30 mm (68%). Our results were similar to those of several previous studies (2,11,17). The lowest rate in nodules larger than 30 mm was probably due to increased vascularity in larger sizes. It might have caused blood-stained materials. Cystic-necrotic degeneration within the larger nodules was probably another factor. Cystic dominance was reported to be an independent finding that increased the possibility of nondiagnostic results, even when experienced operators performed the aspiration (11,18).

Nodule size of less than 5–10 mm was reported to increase the likelihood of nondiagnostic sampling in other studies (11,17). However, in our study, the highest rate was encountered in nodules <10 mm (77.8%), although there was not a statistically significant difference in the adequacy rate among nodules of different sizes. This may be due to the fact that sampling of smaller nodules is technically more difficult, and we were more successful in the sampling of smaller nodules. Nevertheless, it is hard to assume whether nodule size is a significant factor affecting adequacy or not according to our data, because the distribution of nodules according to size was inhomogeneous and sampled nodules in our study were mostly 10–30 mm in diameter.

In addition to these results, we also analyzed the distribution of the adequate samplings according to needle size and nodule size. We found that there was not a statistically significant relation between the needle size and the nodule size in terms of adequacy.

Preparation of slides has been well studied to date, but we have not encountered a study investigating the role of quantity of glass slides. Because overly thick smears might cause inadequacy, the thickness of smears is important (10). The layer thickness of a specimen can be diminished by using more slides for smearing the same amount of sample, and it may also rarefy blood-stained material. Blood-stained material makes microscopic evaluation more difficult, and is more frequently seen in aspiration with thicker needles (2). Furthermore, use of excess slides does not always mean adequate sampling. The greater number of slides, the more time spent by the pathologist for evaluation. The total number of slides is mainly dependent on the amount of sample. One can multiply the quantity in order to obtain thin layered smears. The adequacy rates related to number of slides exhibited variations in our study. Our results revealed that the highest rate of adequate sampling (73.5%) was obtained when five to ten slides were used. Nevertheless, there was not a statistically significant difference between the number of slides and adequacy. On the other hand, we observed that when more slides were used, the rate of adequate samplings increased gradually in procedures performed using 22-G needles, while it decreased gradually in procedures performed using 27-G needles. Our explanation for this is as follows. In the first technique (using 22-G needles), the aspirates were bloody in most of the cases because of the thicker needles. Hence, when the thickness of smears was diminished by using more slides, it resulted in an increased rate of adequate samplings. However, in the second technique (using

27-G needles), increased numbers of slides were caused by multiple punctures. More bleeding into nodules with more punctures resulted in a decreased rate of adequate samplings. Hence, one should try to obtain an optimal number of slides (i.e. five to ten) with a limited number of punctures (one to three). Examining the efficacy of the number of slides with larger series may contribute to the debate on this issue. Although the quantity of slides does not seem to be an absolute factor affecting adequacy according to our results, it is obvious that limitation of the number of slides up to ten is able to yield adequacy. Establishing an upper limit may bring about relief in the workload of pathology departments.

The limitation of this study was that aspiration procedures were performed by using only one size of needle (22-G or 27-G) in the vast majority of the nodules. It would be better if we could perform aspirations with both needles for every nodule.

In conclusion, needle size is an important factor that affects the adequacy of samplings. Aspiration using thinner needles like 27-G may result in less bloody samples increasing the rate. The size of nodules and the number of slides do not affect the adequacy. However, bloody and thicker smears are difficult for pathologists to evaluate, resulting in a decreased rate of adequate sampling.

Acknowledgment

Special thanks should be given to the pathologists, Çiğdem Dicle Arıcan and Serdar Tokmak, for their technical support in this project.

References

- Hoang JK, Lee WK, Lee M, Johnson D, Farrell S. US features of thyroid malignancy: pearls and pitfalls. *Radiographics* 2007; 27: 847-860.
- Degirmenci B, Haktanir A, Albayrak R, Acar M, Sahin DA, Sahin O, Yuçel A, Caliskan G. Sonographically guided fine-needle biopsy of thyroid nodules: the effects of nodule characteristics, sampling technique, and needle size on the adequacy of cytological material. *Clin Radiol* 2007; 62: 798-803.
- Hamburger JI. Diagnosis of thyroid nodules by fine-needle biopsy: use and abuse. *J Clin Endocrinol Metab* 1994; 79: 335-339.
- Gumus M, Cay N, Algin O, Ipek A, Ersoy RU, Belenli O, Ugras S. Comparison of 21 and 27 gauge needles for determining sample adequacy in the aspiration biopsy of thyroid nodules. *Diagn Interv Radiol* 2012; 18: 102-105.
- Tangpricha V, Chen BJ, Swan NC, Sweeney AT, de las Morenas A, Safer JD. Twenty-one-gauge needles provide more cellular samples than twenty-five-gauge needles in fine-needle aspiration biopsy of the thyroid but may not provide increased diagnostic accuracy. *Thyroid* 2001; 11: 973-976.
- Renshaw AA, Pinnar N. Comparison of thyroid fine-needle aspiration and core needle biopsy. *Am J Clin Pathol* 2007; 128: 370-374.
- Gharib H. Fine-needle aspiration biopsy of thyroid nodules: advantages, limitations and effect. *Mayo Clin Proc* 1994; 69: 44-49.
- Goellner JR, Gharib H, Grant CS, Johnson DA. Fine needle aspiration cytology of the thyroid. *Acta Cytol* 1987; 31: 587-590.
- Naim C, Karam R, Edde D. Ultrasound-guided fine-needle aspiration biopsy of the thyroid: methods to decrease the rate of unsatisfactory biopsies in the absence of an on-site pathologist. *Can Assoc Radiol J* 2013; 64: 220-225.
- Cibas ES, Ali SZ. The Bethesda System for Reporting Thyroid Cytopathology. *Am J Clin Pathol* 2009; 132: 658-665.
- Lee YH, Baek JH, Jung SL, Kwak JY, Kim JH, Shin JH. Ultrasound-guided fine needle aspiration of thyroid nodules: a consensus statement by the Korean Society of Thyroid Radiology. *Korean J Radiol* 2015; 16: 391-401.

12. Hamburger JI, Hamburger SW. Fine needle biopsy of thyroid nodules: avoiding the pitfalls. *N Y State J Med* 1986; 86: 241-249.
13. Kim MJ, Kim EK, Park SI, Kim BM, Kwak JY, Kim SJ, Youk JH, Park SH. US-guided fine-needle aspiration of thyroid nodules: indications, techniques, results. *Radiographics* 2008; 28: 1869-1886.
14. de Meer SG, Schreinemakers JM, Zelissen PM, Stapper G, Sie-Go DM, Rinkes IH, Vriens MR. Fine-needle aspiration of thyroid tumors: identifying factors associated with adequacy rate in a large academic center in the Netherlands. *Diagn Cytopathol* 2012; 40: 21-26.
15. Ohori PN, Schoedel KE. Variability in the atypia of undetermined significance/follicular lesion of undetermined significance diagnosis in the Bethesda System for Reporting Thyroid Cytopathology: sources and recommendations. *Acta Cytol* 2011; 55: 492-498.
16. Suen KC. Fine-needle aspiration biopsy of the thyroid. *Can Med Assoc J* 2002; 167: 491-495.
17. Moon HJ, Son E, Kim EK, Yoon JH, Kwak JY. The diagnostic values of ultrasound and ultrasound-guided fine needle aspiration in subcentimeter-sized thyroid nodules. *Ann Surg Oncol* 2012; 19: 52-59.
18. Choi SH, Han KH, Yoon JH, Moon HJ, Son EJ, Youk JH, Youk JH, Kim EK, Kwak JY. Factors affecting inadequate sampling of ultrasound-guided fine-needle aspiration biopsy of thyroid nodules. *Clin Endocrinol* 2011; 74: 776-782.