RESEARCH ARTICLE

Comparison of Pre-Operation Diagnosis of Thyroid Cancer with Fine Needle Aspiration and Core-needle Biopsy: a Meta-analysis

Lei Li[1], Bao-Ding Chen[1], Hai-Feng Zhu[1], Shu Wu[1], Da Wei[2], Jian-Quan Zhang[3], Li Yu[1]*

Abstract

    <u>Background</u>: The aim of this meta-analysis was to compare sensitivities and specificities of fine needle aspiration (FNA) and core needle biopsy (CNB) in the diagnosis of thyroid cancer. <u>Materials and Methods</u>: Articles were screened in Medline, the Cochrane Library, EMBASE and Google Scholar, and subsequently included and excluded based on the patient/problem-intervention-comparison-outcome (PICO) principle. Primary outcome was defined in terms of diagnostic values (sensitivity and specificity) of FNA and CNB for thyroid cancer. Secondary outcome was defined as the accuracy of diagnosis. Compiled FNA and CNB results from the final studies selected as appropriate for meta-analysis were compared with cases for which final pathology diagnoses were available. Statistical analyses were performed for FNA and CNB for all of the selected studies together, and for individual studies using the leave-one-out approach. <u>Results</u>: Article selection and screening yielded five studies for meta-analysis, two of which were prospective and the other three retrospective, for a total of 1,264 patients. Pooled diagnostic sensitivities of FNA and CNB methods were 0.68 and 0.83, respectively, with specificities of 0.93 and 0.94. The areas under the summary ROC curves were 0.905 (±0.030) for FNA and 0.745 (±0.095) for CNB, with no significant difference between the two. No one study had greater influence than any other on the pooled estimates for diagnostic sensitivity and specificity. <u>Conclusions</u>: FNA and CNB do not differ significantly in sensitivity and specificity for diagnosis of thyroid cancer.

Keywords: Thyroid - cancer - diagnosis - fine needle aspiration - core-needle biopsy - ultrasound-guided

Introduction

Thyroid disease is very common with nodules occurring in up to 60% of the general population (Cooper et al., 2009). More than 95% of such nodules are benign, but the incidence of thyroid malignancy has increased more than two-fold worldwide over the last four decades (Sipos and Mazzaferri, 2010). Improved diagnostic capability is a major contributor to the observed increase (Davies and Welch, 2006), but an increase in the true occurrence of thyroid cancer is also partly responsible for the trend (Li et al., 2013). In the United States, 30,000 cases of thyroid cancer are diagnosed annually and about 1,500 people die of the disease (American Cancer Society, 2013). Various types of thyroid malignancies can develop, with varying survival rates, but most such malignancies entail a good prognosis when treatment is appropriate and timely. This makes early, accurate distinction of benign and malignant nodules imperative, since the latter require immediate surgical excision (Yassa et al., 2007) and generally adjuvant radioiodine therapy (RAI) (Morrison et al., 2014).

Often, thyroid malignancy is discovered while at an early stage. In such cases, surgical excision of a nodule deemed suspicious based on clinical examination and history serves a triple purpose. First, analyses via histopathology, cytology, and immunohistopathology of the excised lesion allows for a definitive diagnosis, including the type and grade of the malignancy. Second, histologic and immunochemical examination of margins around the tumor is vital to the surgical staging. Finally, the finding of clear microscopic margins (i.e., no tumor cells in the margins) establishes that an R0 resection has been achieved. When subsequent testing (nuclear medicine and otherwise) confirms no residual tumor, which is particularly common with tumors <1cm, the disease has been effectively cured. Otherwise, the surgical excision constitutes a large component of the treatment, which then can be completed with adjuvant RAI.

Notwithstanding the high accuracy of gold standard pathology tests and the ability to check for clear margins in lesions taken with open biopsy, the fact that less than 5% of thyroid nodules actually are malignant warrants less invasive screening methods. Offering the advantage

[1]Zhenjiang Key Laboratory of Molecular Endocrinology; Affiliated Hospital of Jiangsu University, Zhenjiang, [2]Department of General Surgery, Jiangsu Cancer Hospital, Nanjing, [3]Department of Ultrasound Diagnosis, Shanghai Changzheng Hospital, Shanghai, China *For correspondence: yuli828@163.com

of being completely non-invasive, imaging typically plays a central role in the detection of nodules, especially those that are small, deep, or otherwise not palpable. Often, computed tomography (CT), magnetic resonance imaging (MRI), or positron emission tomography (PET) of the neck for other reasons lead to incidental discovery of a thyroid nodule, but ultrasonography (US) is the imaging modality used most frequency to characterize thyroid nodules whose presence is already known (Nachiappan et al., 2014).

High resolution US has become the gold standard modality for imaging such nodules, because it does not expose patients to ionizing radiation and is particularly revealing in organs close to the surface, such as the thyroid. Standard two-dimensional US can distinguish cystic from solid nodules, as well reveal other aspects of nodular morphology that have implications vis-à-vis the likelihood of malignancy. Microcalcifications, solid composition, and central vascularity all raise the level of suspicion, as do rough edges, a high anteroposterior to transverse diameter (A-P/T), and lack of a halo in the lesion (Alexander et al., 2004; Mansor et al., 2012). Nevertheless, there is a great deal of overlap between features of benign and malignant thyroid nodules using conventional US (Tamsel et al., 2007). It is possible that benign and malignant thyroid nodules can be differentiated much better using US elastography, which involves compressing the target tissue forcefully and analyzing the elastic results (Mansor et al., 2012). However, US elastography may not be particularly revealing when employed as a screening tool for thyroid malignancy in low risk nodular goiter (Vidal-Casariego et al., 2012), the population in which it is most desirable to avoid the invasiveness of biopsy. Cheng et al. (2013) reported an US score based on echogenicity, margins, calcification, and the ratio between the anteroposterior and transversal diameters for predicting whether a lesion is benign or malignant: an US score ≤2 had 80.3% sensitivity and a 72.7% specificity for the lesion being benign.

Weighing the drawback of invasiveness in the case of excision biopsy against the drawback of low accuracy in the case of US leads to an intermediate approach, namely needle biopsy, as an approach to evaluate thyroid nodules. There are two main categories of needle biopsy: Fine needle aspiration (FNA), which supplies a small sample for cytology, and core needle biopsy (CNB), which uses a larger needle to supply more tissue for analysis. FNA is a safe, accurate, and cost-effective method for the initial evaluation of thyroid nodules and is vital in the selection of patients requiring surgical excision, adjuvant treatment, and other clinical interventions (Bukhari et al., 2008; Bongiovanni et al., 2012). The diagnostic specificity and sensitivity of ultrasound-guided FNA are 92% and 83%, respectively (Gharib et al., 2010), but the technique entails certain limitations. Inadequate sampling and indeterminate diagnosis may occur and excisional biopsy may be necessary for definitive diagnosis in many cases (Tandon et al., 2008). Additionally, FNA with cytology (FNAC) of thyroid tissue cannot differentiate between follicular adenoma from follicular carcinoma (Tandon et al., 2008).

To improve the accuracy of needle biopsy of solitary nodules, CNB may be useful, alone or in combination with FNA (Baloch et al., 2008; Park et al., 2011). Minimally invasive and inexpensive compared with open biopsy, CNB constitutes a viable alternative to FNA. Additionally, CNB provides a larger sample than FNA, allowing histopathology and immunohistochemical analysis (Bain et al., 2000; Cheung et al., 2000; Screaton et al., 2003; Ridder et al., 2005). Thus, Hakala et al. (2013) have found that CNB may be more useful than FNA, particularly in the diagnosis of non-follicular thyroid lesions such as papillary carcinoma. Results of another recent study suggest that CNB can reduce non-diagnostic results and the need for surgical diagnosis in patients with calcified thyroid lesions; consequently, the authors conclude that CNB may be preferable to FNA as a first-line diagnostic tool for calcified thyroid nodules (Ha et al., 2014). Additionally, CNB produced accurate, conclusive diagnoses in a study evaluating the technique in patients whose thyroid nodules had been tested previously, but unsuccessfully, via FNA (Yeon et al., 2013).

Controversy surrounds the diagnostic accuracy of CNB. In the setting of thyroid disease, only a few studies have compared the sensitivity and specificity of CNB directly with those of FNA (Quinn et al., 1994; Liu et al., 1995; Karstrup et al., 2001; Harvey et al., 2005; Renshaw and Pinnar, 2007; Sung et al., 2012). Furthermore, notwithstanding the accurate, conclusive diagnoses that are possible when CNB is employed subsequent to failed FNA, this approach entails a selection bias (Screaton et al., 2003; Park et al., 2011; Hahn et al., 2013; Yeon et al., 2013); when used on nodules that have failed to produce diagnostically-useful FNA results, CNB cannot be compared with FNA reliably (Novoa et al., 2012). The meta-analysis reported in this article was conducted for the purpose of comparing the diagnostic sensitivity and specificity of FNA and CNB for thyroid malignancy.

Materials and Methods

Selection of Studies

Articles were selected in literature searches, and subsequently screened in an out based on the patient/problem-intervention-comparison-outcome (PICO) principle. Literature databases searched consisted of Medline, the Cochrane Library, EMBASE, and Google Scholar. The following keywords and key terms were used in the searches: fine needle aspiration/FNA/FNAC; core needle biopsy/coarse needle biopsy/core biopsy/CNB; thyroid nodule; thyroid cancer; diagnostic/diagnosis; ultrasound/US. Articles were selected with publication dates up to December 31, 2013. The resulting lists of references were then hand-searched and studies were selected for further examination, or screened out if they were not relevant.

The remaining articles were screened further for selection of studies to be meta-analyzed. Both prospective and retrospective studies were selected, if they were two-armed with patients undergoing simultaneous FNA and CNB of each nodule, if the needle biopsies were conducted to search for thyroid cancer prior to, or subsequent to, surgical excision, and the needle biopsies were ultrasound-

guided (if conducted prior to surgery) or taken from surgically extracted nodules. Studies were excluded, if only one technique was investigated, or if diagnostic values such as, accuracy, sensitivity and specificity for malignancy were not studied with respect to thyroid nodules. Studies were identified by two independent reviewers using the aforementioned search strategy. When there was uncertainty regarding eligibility, a third reviewer was consulted.

Extraction of Data

From the studies selected for meta-analysis, extracted data consisted of the name of the first author, year of publication, study design, number of patients, patient's age and gender, diagnostic criteria, and the following scientific results: needle gage, needle passes, and number of lesions that were true positives (TP), true negatives (TN), false positives (FP), or false negatives (FN). The accuracy, sensitivity, and specificity or the techniques were also extracted. A positive datum from needle biopsy was categorized as a TP when the final diagnosis on the excised specimen proved to be positive. Similarly, a TN was recorded when a negative result of needle was corroborated by analysis of the surgical specimen. FP was defined as a positive finding on needle biopsy lesions with negative findings on final diagnosis. FN was defined as a negative finding on needle biopsy with a positive finding on final diagnosis.

Quality Assessment and Outcome Measures

The quality of primary studies was evaluated using the Newcastle-Ottawa Scale. This is a validated technique to assess the quality of nonrandomized studies (Wells et al., 2000). Primary outcome was evaluated in terms of diagnostic values (sensitivity and specificity) of FNA and CNB for thyroid cancer, while accuracy of diagnosis was categorized as a secondary outcome.

Statistical Analysis

Data were extracted and analyzed with malignant and suspected malignant lesions defined as positive findings, and benign lesions defined as negative findings. Results of FNA and CNB were compared with the final diagnoses established via histopathology and other gold standard testing to generate numbers for TP, TN, FP, and FN groups. The following values were calculated as follows: Accuracy= (TP+TN)/total, sensitivity=TP/ (TP+FN), specificity=TN/ (TN+FP).

Heterogeneity among the studies was assessed via the Cochran Q and the I^2 statistic. The Q statistic with $p<0.10$ was considered to indicate significant heterogeneity. The I^2 statistic, the percentage of the observed between-study variability caused by heterogeneity, was also considered to indicate heterogeneity when $I^2 >50\%$. When heterogeneity existed between studies, the random-effects model (DerSimonian-Laird approach) was performed. Otherwise, the fixed-effects model (Mantel-Haenszel approach) was employed. The funnel plot for publication bias was not employed in this analysis, because only 5 studies were meta-analyzed, which is inadequate to detect funnel plot asymmetry (Sutton et al., 2000). Difference

between the areas under the curve (AUCs) of two summary receiver operating characteristics (ROC curves) was tested via the Hanley-McNeil method (Hanley and McNeil, 1983; Rosman and Korsten, 2007). Sensitivity analysis for evaluating the influence of each individual study was performed for assessment of sensitivity and specificity of FNA and CNB using the leave-one-out approach. A two-sided $p<0.05$ was considered statistically significant. The homogeneity test (which pooled estimates for sensitivity and specificity), the Moses-Littenberg's approach for summary ROC curves (Moses et al., 1993), and sensitivity analysis (for evaluating the influence of each individual study) were performed using the Meta-Disc version 1.4 (Zamora et al., 2006).

Results

Literature Search

Following the removal of duplicates, the database search identified 112 studies. Screening of these excluded 96 as non-relevant, yielding 16 full-text articles to be assessed for eligibility (Figure 1). Of the 16 studies, one was excluded, because CNB had been performed on nodules in which previous FNA had failed; thus, only CNB had actually been studied. Three studies were excluded, because they analyzed the effects of combined intervention. Another study was excluded, because interventions were not performed on the same nodules. Four studies were excluded because they did not produce outcomes of interest to our study. Two studies were excluded, because the FNA and/or CNB were not ultrasound-guided. This process yielded five studies to be meta-analyzed, which are listed in Table 1 with relevant values.

Characteristics of the Included Studies

Of the five studies in the meta-analysis, two were prospective and the other three were retrospective in design (Table 1). The studies differed greatly in terms of the number of patients, needle gauge for both FNA and CNB, and gender ratio. Across the studies, the number of patients ranged from 52 to 538, the patients tended to be either female or male in majority, and the gauge of needles used ranged from 18-21 for CNB and from 21-25 for FNA. In total, 1264 patients were included, and 859 (68.0%) of them were females.
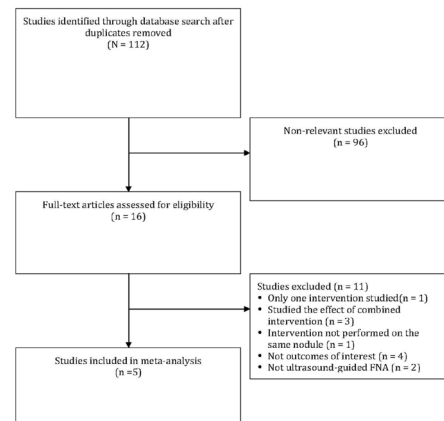

Figure 1. Flow Chart for Study Selection

Table 1. Basic Characteristics of Studies Included in the Meta-Analysis

Study	Study design patients	Number of nodules	Number of nodules	Method	Age, y*	Sex, males (%)	Needle gauge (G) passes	Needle	Newcastle-Ottawa Scale score
Hakala T (2013)	prospective	52	52	FNA	53 ± 17	40 (76.9)	23	2	9
				CNB			20	2	
Na DG (2012)	prospective	220	225	FNA	46 ± 11.7	191 (86.8)	21, 23, 25	1-4	9
				CNB			18	1-3	
Sung JY (2012)	retrospective	538	555	FNA	44.32 ± 11.86	85 (15.8)	21, 23, 25	≥2	9
				CNB			18	1-2	
Renshaw AA (2007)	retrospective	377	n/a	FNA	52 (14-86)	76 (20.2)	21, 23, 25	2-12	9
				CNB			18, 20, 21	1	
Karstrup S (2001)	retrospective	77	n/a	FNA	51(33–81)	13 (16.9)	21	2	9
				CNB			18	2	

n/a, not available; *Data are presented as mean±SD, median (range), or (range)

Table 2. Performance of the Five Diagnostic Studies

Study	Method	Sample size a	TP	TN	FP	FN	Accuracy	Sensitivity	Specificity	Means of final diagnosis
Hakala T (2013)*	FNA	52	12	28	0	12	0.769	0.300	1.000	Surgical resection
	CNB	52	14	27	1	10	0.788	0.583	0.964	
Na DG (2012)*	FNA	149	46	70	0	33	0.779	0.582	1.000	Surgical resection &
	CNB	149	65	69	1	14	0.899	0.823	0.986	Clinical follow-up
Sung JY (2012)*	FNA	555	218	237	0	100	0.820	0.686	1.000	Surgical resection &
	CNB	555	276	235	2	42	0.921	0.868	0.992	Clinical follow-up
Renshaw AA (2007)‡	FNA	62†	29	10	21	2	0.629	0.935	0.323	Surgical resection
	CNB	62†	20	14	17	11	0.548	0.645	0.452	
Karstrup S (2001)§	FNA	40†	15	17	5	3	0.800	0.833	0.773	Surgical resection
	CNB	36†	14	17	1	4	0.861	0.778	0.944	

*Diagnostic criteria for malignancy were Bethesda category 5 and 6. Bethesda category 5 indicates suspicious malignancy and category 6 indicates malignancy; ‡Diagnostic criteria for malignancy were atypical, suspicious, or positive for malignancy; §Diagnostic criteria for malignancy not specified; TP: true positive, the number of cancerous lesions with positive diagnoses; TN: true negative, the number of non-cancerous lesions with negative diagnoses; FP: false positive, the number of non-cancerous lesions with positive diagnoses; FN: false negative, the number of cancerous lesions with negative diagnoses. Accuracy (%)=(TP+TN)/total*100, sensitivity (%)=TP/ (TP+FN)*100%, specificity (%)=TN/(TN+FP)*100%; aThe sample size (total number of nodules, unless otherwise specified) is less in this table compared to Table 1 due to some patients were lost to follow-up and the diagnosis of some nodules was not surgically verified; †indicates the total number of patients
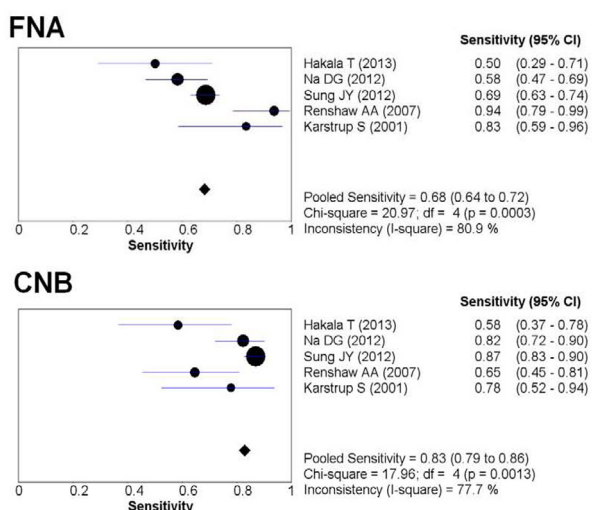

Figure 2. The Forest Plots Showing the Sensitivity of FNA and CNB (Random Effect Approach)

Figure 3. The Forest Plots Showing the Specificity of FNA and CNB (Random Effect Approach)

Comparison of FNA and CNB Diagnostic Performance

The accuracies (proportion of TP + TN) of FNA and CNB for the five studies ranged from 0.629 to 0.820 and 0.548 to 0.921, respectively (Table 2). The sensitivity of FNA and CNB ranged from 0.30 to 0.935 and 0.583 to 0.868, and the specificity ranged from 0.323 to 1.0 and 0.452 to 0.992, respectively.

The random-effects model was used for determining the pooled diagnostic sensitivity since homogeneity tests of sensitivity showed Q=20.97 (p=0.0003) and I²=80.9% for FNA, and Q=17.96 (p=0.0013) and I²=77.7% for CNB.
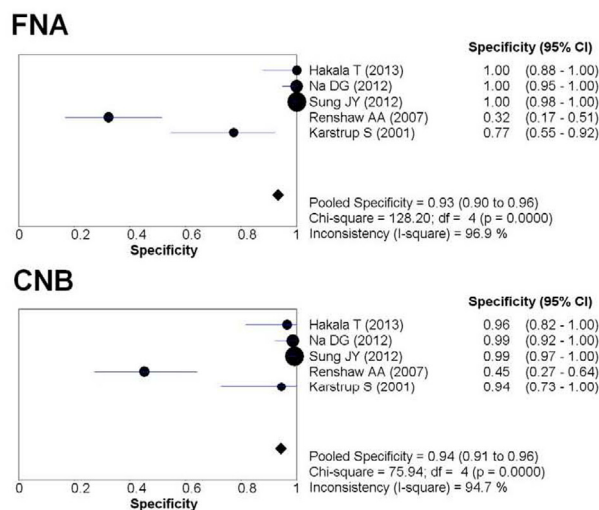
This indicates significant heterogeneity among the studies. The current analysis has revealed pooled diagnostic sensitivities of FNA and CNB methods of 0.68 and 0.83, respectively (Figure 2).

Similarly, the random-effects model was used to evaluate pooled diagnostic specificity as there was also heterogeneity across the studies. The homogeneity tests showed Q=128.20 (p< 0.0001) and I²=96.9% for FNA, and Q=75.94 (p< 0.0001) and I²=94.7% for CNB. The pooled diagnostic specificities of FNA and CNB procedures were 0.93 and 0.94, respectively (Figure 3).

Table 3. The influence of individual studies using the leave-one-out approach

| Method Study excluded	Sensitivity (95% CI)	Specificity (95% CI)
FNA		
Hakala T (2013)	0.691 (0.645, 0.733)	0.928 (0.896, 0.952)
Na DG (2012)	0.701 (0.653, 0.746)	0.918 (0.882, 0.946)
Sung JY (2012)	0.671 (0.590, 0.745)	0.828 (0.758, 0.884)
Renshaw AA (2007)	0.663 (0.617, 0.707)	0.986 (0.968, 0.995)
Karstrup S (2001)	0.675 (0.629, 0.718)	0.943 (0.914, 0.964)
CNB		
Hakala T (2013)	0.841 (0.803, 0.874)	0.941 (0.911, 0.963)
Na DG (2012)	0.829 (0.788, 0.865)	0.933 (0.900, 0.958)
Sung JY (2012)	0.743 (0.666, 0.811)	0.864 (0.798, 0.915)
Renshaw AA (2007)	0.841 (0.803, 0.874)	0.986 (0.967, 0.995)
Karstrup S (2001)	0.830 (0.792, 0.863)	0.943 (0.914, 0.964)
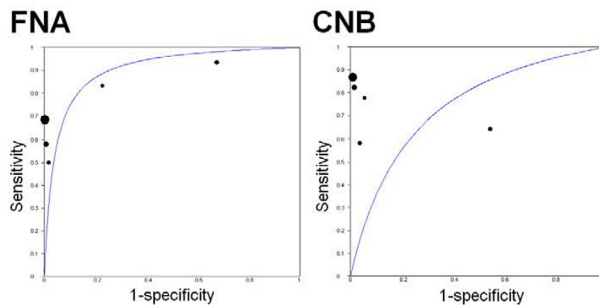
Figure 4. The Summary ROC (sROC) Curves for FNA and CNB. Symmetric sROC Curve Fitted Using Moses' Model. The AUCs for FNA and CNB were 0.905 (with Standard error of 0.030) and 0.745 (with Standard Error of 0.095), Respectively. No Significant Difference was Observed between the two AUCs of FNA and CNB (*p*=0.053).

Summary ROC Curves and Influences of Individual Studies

The areas under the summary ROC curves were 0.905 (standard error=0.030) for FNA and 0.745 (standard error=0.095) for CNB. No significant difference was observed between the summary ROC curves of FNA and CNB (Figure 4). Sensitivity analysis using the leave-one-out approach revealed that no one study had greater influence than any other on the pooled estimates for diagnostic sensitivity and specificity (Table 3).

Discussion

The purpose of this meta-analysis was to comparing FNA and CNB in terms of their diagnostic characteristics as they pertain to the differentiation of malignant and benign thyroid nodules .Despite recent discoveries, this meta-analysis revealed no significant difference in diagnostic values considering both the sensitivity and specificity of FNA and CNB for pre-operation diagnosis for thyroid nodules. No serious complications that needed for hospitalization were registered regarding both techniques in all included studies. For all five studies meta-analyzed, both procedures proved safe with the major adverse events consisting of bleeding and hematomas (Liu et al., 1995; Karstrup et al., 2001; Harvey et al., 2005; Renshaw and Pinnar, 2007; Sung et al., 2012). All studies included in the analysis adopted US-guided biopsies (Hakala et al., 2013), except when biopsies were

taken during intraoperative surgical specimen. All studies included in this analysis compared the two techniques on the same nodules, thus allowing direct and more objective comparison that many of the studies that were eliminated on the article screening process.

Two meta-analyses by Tandon et al. (2008) and the by Novoa et al. (2012) evaluated FNA and CNB in diagnosing malignancy of the head and neck. Tandon et al. (2008) included 12 studies and elucidated a FNA sensitivity of 90.7% and a specificity of 98.1% for detecting malignancy, higher than an sensitivity of 69% for detecting recurrent cancer. Novoa et al. (2012) included 16 studies assessing CNB and determined that CNB could detect the malignancy with a sensitivity 80% and specificity of 100%. In both meta-analyses, the low sensitivity detected a high number of FPs, which is less of a problem than a high number of FNs.

In the five studies included in the current meta-analysis, the FP rate ranged from 0% to 5.8% for FNA, and 0.4% to 27% for CNB. The range of FPs across the studies may reflect differences in accuracy of the cytopathology and histophathology among the studies. Consistent with the results of the current meta-analysis, comparison of the previous CNB and FNA meta-analyses revealed no significant difference in sensitivities of CNB and FNA (p=0.350) (Novoa et al., 2012). In contrast, the results of both the Tandon et al. (2008) and suggest that CNB is more accurate and specific, with a better negative predictive value compared with FNA when applied to neoplasia throughout the head and neck. However, neither study compared FNA and CNB in terms of their diagnostic values specifically with respect to the thyroid.

The results of the current study should be interpreted in the context of several limitations. There was bias with respect to sampling, firstly because all studies tended to focus on one gender. As a whole, female (68%) were included as the majority, which may overlook some underlying bias. Secondly, in one study patients were included based on prior preoperative FNA diagnoses that warranted surgical treatment (Hakala et al., 2013). In another study (Na et al., 2012), however, patients were included due to previous unsuccessful FNA results; this may suggest lesions that were difficult to aspirate. Biopsy techniques might be beneficial to different subtype of lesions in patients (Hahn et al., 2013; Hakala et al., 2013), and such heterogeneous sampling might have obscured the analysis. For the purpose of the current study, the calculations were based on the assumption that only malignant or possibly malignant samples constitute positive findings, and the results were compared with the 'gold standard' testing of excised nodules. Consequently, this analysis may be limited in terms of comparison of the sensitivity, specificity, and accuracy of the techniques. Other aspects, such as rate of inconclusive diagnoses (e.g., non-diagnostic and AUS/FLUS) were not considered. The current meta-analysis also excluded biopsies with insufficient material for diagnostic evaluation. Additionally, there was heterogeneity with respect to diagnostic criteria: the Bethesda system for reporting thyroid cytopathology was used for interpretation of

biopsies in only three studies of the five studies (Na et al., 2012; Sung et al., 2012; Hakala et al., 2013). Furthermore, diagnostic criteria for malignancy were compared with surgical resections in three studies (Karstrup et al., 2001; Renshaw and Pinnar, 2007; Hakala et al., 2013) as the final diagnosis. However, in the other two studies (Na et al., 2012; Sung et al., 2012) final diagnoses were determined by histopathological results after surgical resection (for some benign and all malignant nodules), as well as by clinical follow-ups (for benign nodules). Future meta-analyses should seek to eliminate studies with such heterogeneity.

Due to the outcome of the literature selection process, the number of studies included in the meta-analysis was small, and of the five studies included, only two were prospective in design. Additionally, in none of the five studies was the pathologist blinded to the results of the FNA and CNB when determining the final diagnosis. The studies also differed in the diagnostic categories used and the definition of malignancy. Notably, the term "malignant" in two of the studies included "suspected malignancy" (Renshaw and Pinnar, 2007; Sung et al., 2012). Moreover, the analysis included multiple forms of thyroid cancer, yet it is possible that CNB and FNA may have better diagnostic abilities in identifying specific types of thyroid malignancy. These limitations reflect the very small pool of available studies; therefore, an additional conclusion of the current study is that more prospective studies comparing FNA and CNB cytology and gold standard diagnosis of excised nodules in thyroid need to be performed.

In summary, the current study suggests that both FNA and CNB are simple, minimally invasive, and safe methods for reliable diagnosis of thyroid malignancy. The fact that both procedures show a reasonable number of FPs with respect to the detection of thyroid cancer suggests that that they should not replace excisional biopsy, at least not for certain patients (Novoa et al., 2012). Additionally, as noted above, the revelation of just a handful of appropriate studies by the literature search and article screening process highlights the need for well-designed prospective studies to investigate this issue.

Acknowledgements

References

Alexander EK, Marqusee E, Orcutt J, et al (2004). Thyroid nodule shape and prediction of malignancy. *Thyroid*, **14**, 953-8.

Bain G, Bearcroft PW, Berman LH, Grant JW (2000). The use of ultrasound-guided cutting-needle biopsy in paediatric neck masses. *Eur Radiol*, **10**, 512-5.

Baloch ZW, Cibas ES, Clark DP, et al (2008). The national cancer institute thyroid fine needle aspiration state of the science conference: A summation. *Cytojournal*, **5**, 6.

Bongiovanni M, Spitale A, Faquin WC, Mazzucchelli L, Baloch ZW (2012). The bethesda system for reporting thyroid cytopathology: A meta-analysis. *Acta Cytol*, **56**, 333-9.

Bukhari MH, Niazi S, Hanif G, et al (2008). An updated audit of fine needle aspiration cytology procedure of solitary thyroid nodule. *Diagn Cytopathol*, **36**, 104-12.

Cheng P, Chen ED, Zheng HM, He QX, Li Q (2013). Ultrasound score to select subcentimeter-sized thyroid nodules requiring ultrasound-guided fine needle aspiration biopsy in eastern China. *Asian Pac J Cancer Prev*, **14**, 4689-92.

Cheung YC, Wan YL, Lui KW, Lee KF (2000). Sonographically guided core-needle biopsy in the diagnosis of superficial lymphadenopathy. *J Clin Ultrasound*, **28**, 283-9.

Cooper DS, Doherty GM, Haugen BR, et al (2009). Revised american thyroid association management guidelines for patients with thyroid nodules and differentiated thyroid cancer. *Thyroid*, **19**, 1167-214.

Davies L, Welch HG (2006). Increasing incidence of thyroid cancer in the united states, 1973-2002. *JAMA*, **295**, 2164-7.

Gharib H, Papini E, Paschke R, et al (2010). American association of clinical endocrinologists, associazione medici endocrinologi, and europeanthyroid association medical guidelines for clinical practice for the diagnosis and management of thyroid nodules. *Endocr Pract*, **16**, 1-43.

Ha EJ, Baek JH, Lee JH, et al (2014). Core needle biopsy can minimise the non-diagnostic results and need for diagnostic surgery in patients with calcified thyroid nodules. *Eur Radiol*, **24**, 1403-9.

Hahn SY, Shin JH, Han BK, Ko EY, Ko ES (2013). Ultrasonography-guided core needle biopsy for the thyroid nodule: Does the procedure hold any benefit for the diagnosis when fine-needle aspiration cytology analysis shows inconclusive results? *Br J Radiol*, **86**, 20130007.

Hakala T, Kholova I, Sand J, Saaristo R, Kellokumpu-Lehtinen P (2013). A core needle biopsy provides more malignancy-specific results than fine-needle aspiration biopsy in thyroid nodules suspicious for malignancy. *J Clin Pathol*, **66**, 1046-50.

Hanley JA, McNeil BJ (1983). A method of comparing the areas under receiver operating characteristic curves derived from the same cases. *Radiology*, **148**, 839-43.

Harvey JN, Parker D, De P, Shrimali RK, Otter M (2005). Sonographically guided core biopsy in the assessment of thyroid nodules. *J Clin Ultrasound*, **33**, 57-62.

Karstrup S, Balslev E, Juul N, Eskildsen PC, Baumbach L (2001). Us-guided fine needle aspiration versus coarse needle biopsy of thyroid nodules. *Eur J Ultrasound*, **13**, 1-5.

Li N, Du XL, Reitzel LR, Xu L, Sturgis EM (2013). Impact of enhanced detection on the increase in thyroid cancer incidence in the united states: Review of incidence trends by socioeconomic status within the surveillance, epidemiology, and end results registry, 1980-2008. *Thyroid*, **23**, 103-10.

Liu Q, Castelli M, Gattuso P, Prinz RA (1995). Simultaneous fine-needle aspiration and core-needle biopsy of thyroid nodules. *Am Surg*, **61**, 628-32.

Mansor M, Okasha H, Esmat S, et al (2012). Role of ultrasound elastography in prediction of malignancy in thyroid nodules. *Endocr Res*, **37**, 67-77.

Morrison SA, Suh H, Hodin RA (2014). The surgical management of thyroid cancer. *Rambam Maimonides Med J*, **5**, 8.

Moses LE, Shapiro D, Littenberg B (1993). Combining independent studies of a diagnostic test into a summary roc curve: Data-analytic approaches and some additional considerations. *Stat Med*, **12**, 1293-316.

Na DG, Kim JH, Sung JY, et al (2012). Core-needle biopsy is

more useful than repeat fine-needle aspiration in thyroid nodules read as nondiagnostic or atypia of undetermined significance by the bethesda system for reporting thyroid cytopathology. *Thyroid*, **22**, 468-75.

Nachiappan AC, Metwalli ZA, Hailey BS, et al (2014). The thyroid: Review of imaging features and biopsy techniques with radiologic-pathologic correlation. *Radiographics*, **34**, 276-93.

Novoa E, Gurtler N, Arnoux A, Kraft M (2012). Role of ultrasound-guided core-needle biopsy in the assessment of head and neck lesions: A meta-analysis and systematic review of the literature. *Head Neck*, **34**, 1497-503.

Park KT, Ahn SH, Mo JH, et al (2011). Role of core needle biopsy and ultrasonographic finding in management of indeterminate thyroid nodules. *Head Neck*, **33**, 160-5.

Quinn SF, Nelson HA, Demlow TA (1994). Thyroid biopsies: Fine-needle aspiration biopsy versus spring-activated core biopsy needle in 102 patients. J Vasc Interv Radiol, 5, 619-23.

Renshaw AA, Pinnar N (2007). Comparison of thyroid fine-needle aspiration and core needle biopsy. *Am J Clin Pathol*, **128**, 370-4.

Ridder GJ, Technau-Ihling K, Boedeker CC (2005). Ultrasound-guided cutting needle biopsy in the diagnosis of head and neck masses. *Laryngoscope*, **115**, 376-7.

Rosman AS, Korsten MA (2007). Application of summary receiver operating characteristics (sroc) analysis to diagnostic clinical testing. *Adv Med Sci*, **52**, 76-82.

Screaton NJ, Berman LH, Grant JW (2003). Us-guided core-needle biopsy of the thyroid gland. *Radiology*, **226**, 827-32.

Sipos JA, Mazzaferri EL (2010). Thyroid cancer epidemiology and prognostic variables. *Clin Oncol (R Coll Radiol)*, **22**, 395-404.

American Cancer Society (2013). Cancer facts & figures 2013. Available at: http://www.cancer.org/acs/groups/content/@epidemiologysurveilance/documents/document/acspc-036845.pdf. Accessed August 21, 2014.

Sung JY, Na DG, Kim KS, et al (2012). Diagnostic accuracy of fine-needle aspiration versus core-needle biopsy for the diagnosis of thyroid malignancy in a clinical cohort. *Eur Radiol*, **22**, 1564-72.

Sutton AJ, Duval SJ, Tweedie RL, Abrams KR, Jones DR (2000). Empirical assessment of effect of publication bias on meta-analyses. *BMJ*, **320**, 1574-7.

Tamsel S, Demirpolat G, Erdogan M, et al (2007). Power doppler us patterns of vascularity and spectral doppler us parameters in predicting malignancy in thyroid nodules. *Clin Radiol*, **62**, 245-51.

Tandon S, Shahab R, Benton JI, et al (2008). Fine-needle aspiration cytology in a regional head and neck cancer center: Comparison with a systematic review and meta-analysis. *Head Neck*, **30**, 1246-52.

Vidal-Casariego A, Lopez-Gonzalez L, Jimenez-Perez A, et al (2012). Accuracy of ultrasound elastography in the diagnosis of thyroid cancer in a low-risk population. *Exp Clin Endocrinol Diabetes*, **120**, 635-8.

Wells GA, Shea B, O'Connell D, et al (2000). The newcastle-ottawa scale (nos) for assessing the quality of nonrandomized studies in meta-analyses. The 3rd symposium on systematic reviews: Beyond the basics. Oxford, England.

Yassa L, Cibas ES, Benson CB, et al (2007). Long-term assessment of a multidisciplinary approach to thyroid nodule diagnostic evaluation. *Cancer*, **111**, 508-16.

Yeon JS, Baek JH, Lim HK, et al (2013). Thyroid nodules with initially nondiagnostic cytologic results: The role of core-needle biopsy. *Radiology*, **268**, 274-80.

Zamora J, Abraira V, Muriel A, Khan K, Coomarasamy A (2006). Meta-disc: A software for meta-analysis of test accuracy data. *BMC Med Res Methodol*, **6**, 31.