RESEARCH ARTICLE

Diagnostic Accuracy of Ultrasonography in Differentiating Benign and Malignant Thyroid Nodules Using Fine Needle **Aspiration Cytology as the Reference Standard**

Tariq Alam, Yasir Jamil Khattak, Madiha Beg*, Abdul Raouf, Muhammad Azeemuddin, Asif Alam Khan

Abstract

Background: In Pakistan thyroid cancer is responsible for 1.2% cases of all malignant tumors. Ultrasonography (US) is helpful in detecting cancerous thyroid nodules on basis of different features like echogenicity, margins, microcalcifications, size, shape and abnormal neck lymph nodes. We therefore aimed to calculate diagnostic accuracy of ultrasound in detection of carcinoma in thyroid nodules taking fine needle aspiration cytology as the reference standard. Materials and Methods: A cross-sectional analytical study was designed to prospectively collect data from December 2010 till December 2012 from the Department of Radiology in Aga Khan University Hospital, Karachi, Pakistan. A total of 100 patients of both genders were enrolled after informed consent via applying non-probability consecutive sampling technique. Patients referred to Radiology department of Aga Khan University to perform thyroid ultrasound followed by fine-needle aspiration cytology of thyroid nodules were included. They were excluded if proven for thyroid malignancy or if their US or FNAC was conducted outside our institution. Results: The subjects comprised 76 (76%) females and 24 males. Mean age was 41.8±SD 12.3 years. Sensitivity and specificity with 95%CI of ultrasound in differentiating malignant thyroid nodule from benign thyroid nodule calculated to be 91.7% (95%CI, 0.72-0.98) and 78.94% (0.68-0.87) respectively. Reported positive predictive value and negative PV were 57.9% (0.41-0.73) and 96.8% (0.88-0.99) and overall accuracy was 82%. Likelihood ratio (LR) positive was computed to be 4.3 and LR negative was 0.1. Conclusions: Ultrasonography has a high diagnostic accuracy in detecting malignancy in thyroid nodules on the basis of features like echogenicity, margins, micro calcifications and shape.

Keywords: Cancer of thyroid - ultrasound - fine needle aspiration - sensitivity and specificity - Pakistan

Asian Pac J Cancer Prev, 15 (22), 10039-10043

Introduction

Thyroid nodules are one of the most common presentations of many benign and malignant thyroid diseases (Cooper et al., 2006; Yumei et al., 2011). Thyroid disease is a major public health problem in Pakistan because of iodine deficient diet and according to United Nations International Children's Emergency Fund (UNICEF), 70% of total population in Pakistan is at risk of developing thyroid diseases due to iodine deficiency (Akhtar et al., 2004; Sushel et al., 2009). In Pakistan, thyroid cancer is responsible for 1.2% cases of all malignant tumors (Zuberi et al., 2004). According to Surveillance Epidemiology and End Results (SEER) and Cancer Statistic Review 1975-2007, the prevalence of thyroid cancer is 0.08% (Altekruse et al., 2007).

It has been estimated that about 5-10% of all thyroid nodules are malignant (Welker et al., 2003). Therefore, imaging importance of thyroid nodules rests with the need to exclude thyroid cancer by various imaging modalities.

Although a number of imaging modalities like radionuclide thyroid scan, computed tomography and magnetic resonance imaging are used for diagnosis of thyroid diseases, but Ultrasonography (US) is a frequent technique, especially for evaluation of thyroid nodules because of its cost effectiveness, safety and noninvasiveness (Koike et al., 2001).

US is helpful in detecting cancer in thyroid nodules on basis of different features like echogenicity, margins, micro calcifications, size, shape, internal contents and abnormal neck lymph nodes (Moon et al., 2008; Ahn et al., 2010; Cheng et al., 2013). In a study conducted by Kim et al out of 49 fine needle aspiration biopsy proven malignant nodules, US detected 46 nodules constituting 93.8% of the total (Kim et al., 2002).

To the best of our knowledge, documented evidence was deficient from the sub-continent, specifically from Pakistani population, which aimed to evaluate the ability

Department of Radiology, The Aga Khan University Hospital, Karachi Pakistan *For correspondence: madiha.beg@aku.edu; drmadiha@gmail.com

of US to differentiate between benign and malignant thyroid nodules and its association with fine-needle aspiration cytology (FNAC). Consequently, on account of thyroid disease being a major health problem in our country (Akhtar et al., 2004; Zuberi et al., 2004; Sushel et al., 2009), a study evaluating accuracy of US in detecting malignancy in thyroid nodules was deemed essentially necessary in the Pakistani population. Therefore, we aimed to compute diagnostic accuracy with 95% confidence intervals of ultrasound in detection of carcinoma in thyroid nodules taking Fine Needle Aspiration Cytology as reference standard

Materials and Methods

Overview

A cross-sectional analytical study was designed to prospectively collect data from the department of Radiology, Aga Khan University Hospital (AKUH) Karachi from December 2010 to December 2012 via non-probability consecutive sampling technique. This original research is based on a thesis; the study protocol was a priori approved by the Research Evaluation Unit of The College of Physicians and Surgeons Pakistan.

Study population

Patients of either gender presenting with palpable thyroid nodule (diagnosed by primary physician on clinical examination) referred to Radiology department of AKUH for thyroid ultrasound followed by fine-needle aspiration cytology of thyroid nodules were included after an informed consent. They were excluded if proven for thyroid malignancy or if their US or FNAC was conducted outside our institution.

US Image analysis

A thyroid nodule was considered positive on Ultrasonography (US) or malignant if one or more than one of following sonographic features were found: micro calcification defined as punctuate (less than 2mm) hyper echoic foci either with or without acoustic shadows (Kim et al., 2002). Micro-lobulation was characterized as presence of many small lobules on surface of a nodule or irregular margins (Kim et al., 2002). Marked hypo echogenicity demarcated as decreased echogenicity compared with surrounding neck muscle (Kim et al., 2002). Shape characterized as shape of nodule taller than wider (Kim et al., 2002).

Thyroid nodule was categorized as negative, that is, malignancy not found if none of the above sonographer feature was seen.

A thyroid nodule was considered True positive (TP), i.e. malignancy detected, when specimen showed cytological atypical cells (nuclei crowded and overlapping, enlarged and pleomorphic and a thyroid nodule was considered True negative (TN)., that is, malignancy not established, when specimen did not exhibit cytological atypical cells (nuclei crowded and overlapping, enlarged and pleomorphic) (Moon et al., 2008; Ahn et al., 2010).

A thyroid nodule was considered false positive (FP) when ultrasound findings were suggestive of malignancy

however specimen did not show cytological atypical cells. Similarly a thyroid nodule was considered false negative (FN), when it was reported as benign on ultrasound findings however the histopathology was suggestive of malignancy.

Positive predictive values (PPV) were calculated to determine the probability of a patient having thyroid malignancy when the nodule was reported as malignant on ultrasonography. Similarly negative predictive values (NPV) were calculated to find the proportion of patients, not having malignancy amongst all those which were reported as having benign on ultrasonography.

Diagnostic accuracy was calculated to express the proportion of correctly classified subjects (Patients having thyroid malignancy+patients with benign thyroid nodules, TP+TN) among all subjects (TP+TN+FP+FN).

Nodule size was not included in the criteria as it is not predictive of malignancy, because the likelihood of cancer in a thyroid nodule has been shown to be the same regardless of the size measured at US (Kim et al., 2002).

FNAC analysis

All Ultrasounds were followed by FNAC by a single consultant radiologist with more than 5 years' of experience in performing the procedure. Cytopathologist with 5 years' of experience analyzed the FNAC specimen who was blinded to US diagnosis of thyroid nodule in order to control bias. FNAC diagnosis of malignancy in thyroid nodules was acquired from medical record system and was taken as reference standard for this study.

Ultrasound imaging protocol

All of the ultrasounds of thyroid gland were performed by a single radiologist on Nemio XG ultrasound machine equipped with 3.5-5MHz Curvilinear and 7.5-15 MHz Linear probe. Both transverse and longitudinal images were taken and send to the Picture and Archiving System

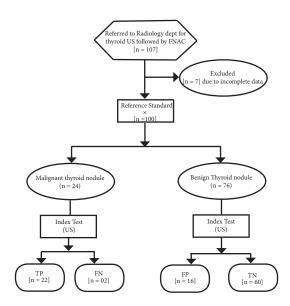


Figure 1. Flow Chart for Study assessing Diagnostic Accuracy of Ultrasonography in Differentiating Thyroid Nodules by Utilizing FNAC as Reference Standard

(PACS) for latter review as well. Sample size calculation

Reported sensitivity of ultrasound for detection of thyroid carcinoma is 93.8% (Yunus et al., 2010), taking confidence interval of 95%, margin of error 5%, the calculated sample size N = 90.

$$n = \frac{Z_{1-\alpha/2}^2 - P(1-P)}{B^2}$$

Data analysis procedure

Data were entered and analyzed in SPSS 20.0 version. Descriptive analyses were run to compute frequencies and percentages for categorical variables like sex, malignancy in thyroid nodules. Mean and standard deviation computed for quantitative variable like age. Sensitivity, Specificity, Positive predictive value (PPV), Negative predictive value (NPV) with 95% Confidence Intervals (CI) were computed for individual ultrasound characteristic feature for malignancy of thyroid nodule. In addition, diagnostic accuracy was also reported of each feature. Diagnostic accuracy of ultrasound in detection of malignancy in thyroid nodules was also calculated with FNAC as reference standard with 95%CI. Likelihood ratios were also computed and reported.

Results

The sample comprised of 76 (76%) female out of a total of 107. Male to female ratio was 1:3. Mean age was 41.77±SD 12.31 years. Age range was 17-72 years. A total of 7 patients were excluded due to incomplete data. From the sample, our study results depicted overall sensitivity, specificity with 95%CI of ultrasound in differentiating malignant thyroid nodule from benign thyroid nodule to be 91.66% [95%CI (0.72, 0.98)] and 78.94% [95%CI (0.68,0.87)] respectively. Reported PPV and NPV were 57.89% [95%CI (0.41,0.73) and 96.77% [95%CI (0.88, 0.99)]. We computed overall accuracy as 82%. The ultrasound features of malignant thyroid nodules observed in our study were marked hypo-echogenicity,

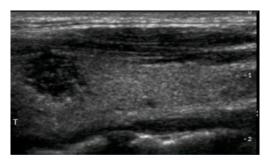


Figure 2. US of Left Lobe of Thyroid Longitudinal View Showing A Markedly Hypo Echoic Nodule with **Irregular Margins and Micro Calcifications**

micro lobulated/irregular margins, and shape "taller than wide" and micro calcifications. True positive rate (TPR) for individual sonographic features was also calculated. TPR for marked hypo-echogenicity was 10/24=41.67%, for micro lobular/irregular margins TPR: 3/24=12.5%. Thyroid nodules having shape taller than wide, that is, vertically oriented was observed in 10/24=41.66% and Micro-calcification observed as 8/24=33.33%. Likelihood ratio (LR) positive was computed to be 4.3 and likelihood ratio negative was 0.1.

Table 1 shows individual sonographic features computed sensitivity, specificity, PPV, NPV and accuracy with 95%CI.

Discussion

Thyroid nodules are very common (Unnikrishnan AG et al., 2011; Taddesse A et al., 2011). They are found in 4%-8% of adults by means of palpation, in 10%-41% by means of US (Carroll et al., 1982; Brander et al., 1991; Bruneton et al., 1994; Wiest et al., 1998; Nam-Goong et al., 2004; Rahimi et al., 2013) and in 50% by means of pathologic examination at autopsy. The prevalence of thyroid nodules increases with age. The likelihood that a nodule is malignant is affected by a variety of risk factors. It has been estimated that about 5-10 % of all thyroid nodules are malignant (Welker et al., 2003). Therefore imaging importance of thyroid nodules rests with the need to exclude thyroid cancer by various imaging modalities.

Although a number of imaging modalities like radionuclide thyroid scan, computed tomography and magnetic resonance imaging are used for the diagnosis of thyroid diseases, but US is a commonly employed technique, especially for evaluation of thyroid nodules (Shin et al., 2014). It is mainly because of its cost effectiveness, safety and noninvasiveness (Koike et al., 2001).

Many studies have been published in which the ability to predict whether a thyroid nodule is benign or malignant on the basis of US findings was assessed. Nodule size is

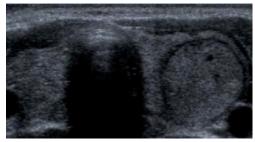


Figure 3. US of Thyroid Transverse View Showing An Iso-Echoic Nodule in Left Lobe of Thyroid Gland with Well-defined Margins which Proved to be a Benign **Hyper Plastic Nodule on FNAC**

Table 1. Accuracy of Individual Ultrasonographic Features/Characterstics for Malignant Thyroid Nochules

Characteristics of US features	Sensitivity (95% CI*)	Specificity (95% CI*)	^ PPV (95% CI*)	¤ NPV (95% CI*)	Accuracy
Hypoechogenicity	41.66% (0.23, 0.63)	85.52% (0.75, 0.92)	47.62% (0.26, 0.69)	82.27% (0.72, 0.89)	75% (75/100)
Microlobulated	12.5% (0.03, 0.33)	94.74% (0.86, 0.98)	42.85% (0.12, 0.79)	77.42% (0.67, 0.85)	75% (75/100)
Vertical Shape	41.66% (0.23, 0.63)	97.37% (0.89, 0.99)	83.33% (0.51, 0.97)	83.33% (0.51, 0.97)	84% (84/100)
Micro Calcification	33.33% (0.16, 0.55)	93.42% (0.85, 0.97)	61.54% (0.32, 0.84)	81.61% (0.72, 0.88)	79% (79/100)

^{*}CI -Confidence Interval; ^ PPV-Positive Predictive value; ¤ NPV-Negative predictive value

not predictive of malignancy, because the likelihood of cancer in a thyroid nodule has been shown to be the same regardless of the size measured at US (Solbiati et al., 2001; Papini et al., 2002; Frates et al., 2004).

Several US features have been found to be associated with an increased risk of thyroid cancer, including presence of calcifications, hypo echogenicity, irregular margins, and absence of a halo, predominantly solid composition, and intranodule vascularity. However, the sensitivities, specificities, and negative and positive predictive values for these criteria are extremely variable from study to study, and no US feature has both a high sensitivity and a high positive predictive value for thyroid cancer. The combination of factors improves the positive predictive value of US to some extent (Frates et al., 2004).

In our study, the shape of a nodule 'taller than wider' has the highest sensitivity, specificity, PPV, NPV and Diagnostic accuracy (95%CI) of 41.66% [95%CI(0.23,0.63)], 97.37% [95%CI(0.89,0.99)], 83.33% [95%CI(0.51,0.97)], 84.09% [95%CI(0.74,0.91)] and 84% respectively which support previous study results (Kim et al., 2002; Alexander et al., 2004) that suggest a taller than-wide shape is very specific for differentiating malignant thyroid nodules from benign ones. This result conveys the fact that malignant nodules (taller than wide) grow across normal tissue planes, while benign nodules grow parallel to normal tissue planes (Papini et al., 2002; Alexander et al., 2004).

Our study results depicted micro lobulated/irregular margin of a nodule have sensitivity, specificity, PPV,NPV and diagnostic accuracy (95%CI) of 12.5% [95%CI(0.03,0.33)], 94.74% [95%CI(0.86,0.98)], 42.85% [95%CI(0.12,0.79)], 77.42% [95%CI(0.67,0.85)] and 75% respectively. Results of earlier reports (Alexander et al., 2004; Frates et al., 2004) have suggested that blurred or ill-defined nodular margins favor a diagnosis of malignancy.

However, with the development of high-frequency transducer US techniques, a previously described ill-defined margin could actually be a micro-lobulated / irregular margin with sharp demarcation or a poorly defined margin in which the tumor cannot be discriminated from the normal parenchyma. When the marginal tumor infiltration of malignancy is minimal, it can be seen as an ill-defined margin. Therefore, we hypothesized that an irregular margin is suggestive of malignancy, while an ill-defined margin can be seen in both benign and malignant nodules. This hypothesis is further reinforced by our findings.

We found association of markedly hypo echogenicity was significantly different between benign and malignant nodules (p<0.0001) by Fischer exact test. Most of malignant thyroid nodule was markedly hypo echogenic in comparison to the benign nodules, as obvious from Figures 2 and 3. Stated sensitivity, specificity, PPV, NPV and diagnostic accuracy (95%CI) of marked hypo echogenicity was 41.66% [95%CI(0.23,0.63)], 85.52% [95%CI(0.75,0.92)], 47.62% [95%CI(0.26,0.69)], 82.27% [95%CI(0.72,0.89)] and accuracy 75% respectively. These findings are in accordance with previous studies (Kim et al., 2002; Papini et al., 2002; Wienke et al., 2003) that were conducted on a relatively smaller scale.

We found that micro calcifications have a high specificity of 93.42% but a low sensitivity of 33.33% which is in accordance to previous studies (Kim et al., 2002; Papini et al., 2002). Microcalcifications correspond pathologically to calcified psammoma bodies that are typical of papillary cancer (Meissner et al., 1958; Klinck et al., 1959; Pusztaszeri et al., 2013). Macro calcification or coarse calcifications are related to fibrosis and degeneration (Wang et al., 2006). Benign nodules have coarse calcifications, especially with long disease duration (Komolafe et al., 1981; Kuma et al., 1992; Vinayak S et al 2012). In this study, the overall sensitivity of thyroid US for depicting a malignant nodule is reported as 82%.

Our study had some limitations. The first limitation is that single radiologist performed all the ultrasounds, it would have been better if more than one radiologist was involved and then we would have been successful in calculating the inter-observer variability. Secondly, the small number of the malignant nodules was another limitation. But this inherently minimized reporting bias in our study. However, the reported malignancy in thyroid nodules in the existing literature is 5-10% (Welker et al., 2003). We executed a single-center study; results of a multi-center study would be more generalizable. We have taken FNAC as reference standard but it would have been better if surgical biopsy was taken instead. Confidence intervals for sensitivity, specificity, PPV and NPV have not been earlier calculated and reported from our area. Other important features like internal vascularity and presence or absence of cervical lymph nodes can be also evaluated which may further increase the sensitivity and specificity of US in detection of malignancy in thyroid nodules. This is strength of our study as estimates of diagnostic accuracy and measures of statistical uncertainty are reported. We did not exclude any patient on the basis of negative index test ; neither were they deprived to be tested for the reference standard. Thereby, minimizing partial verification bias.

Also, results of the index test were interpreted without knowledge of results of the reference standard.

However, we did not take into account the indeterminate or intermediate test results.

In conclusion, US have a high diagnostic accuracy in detecting malignancy in thyroid nodules on basis of features like echogenicity, margins, micro calcifications and shape. Radiologists' must be familiar with these signs on ultrasound that aid to distinguish benign from malignant thyroid nodules and thus avoiding unnecessary FNAC.

References

Ahn SS, Kim EK, Kang DR, et al (2010). Biopsy of thyroid nodules: comparison of three sets of guideline. *Am J Roentgenol*, **194**, 31-7.

Akhtar T, Zahoor Ullah, Paracha PI, Lutfullah G (2004). Impact assessment of salt iodization on the prevalence of goiter in district Swat. *Pak J Med Sci*, **20**, 303-7.

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

Altekruse SF, Kosary CL, Krapcho M, et al (2007). SEER cancer statistics review. *Nat Cancer Inst*. [Online]. URL:http://seer.

- cancer.gov/csr.
- Brander A, Viikinkoski P, Nickels J, Kivisaari L(1991). Thyroid gland: US screening in a random adult population. Radiology, 181, 683-7.
- Bruneton JN, Balu-Maestro C, Marcy PY, Melia P, Mourou MY (1994). Very high frequency (13 MHz) ultrasonographic examination of the normal neck: detection of normal lymph nodes and thyroid nodules. J Ultrasound Med, 13, 87-90.
- Carroll BA (1982). Asymptomatic thyroid nodules: incidental sonographic detection. Am J Roentgenol, 138, 499-501.
- Cheng P, Chen ED, Zheng HM, He QX, Li Q, et al (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.
- Cooper DS, Doherty GM, Haugen BR, et al, (2006). American thyroid association guidelines taskforce. management guidelines for patients with thyroid nodules and differentiated thyroid cancer. Thyroid, 16, 109-42.
- Frates MC, Benson CB, Doubilet PM (2004). Likelihood of thyroid cancer based on sonographic assessment of nodule size and composition [abstract]. in: radiological society of North America Scientific assembly and annual meeting program. oak brook, III: Radiological Society of North America, 395.
- Kim EK, Park CS, Chung WY, et al (2002). New sonographic criteria for recommending fine-needle aspiration biopsy of nonpalpable solid nodules of the thyroid. Am J Roentgenol, **178**, 687-91.
- Klinck GH, Winship T (1959). Psammoma bodies and thyroid cancer. Cancer, 12, 656-62.
- Koike E, Noguchi S, Yamashita H, et al (2001). Ultrasonographic characteristics of thyroid nodules: prediction of malignancy. Arch Surg, 136, 334-7.
- Komolafe F (1981). Radiological patterns and significance of thyroid calcification. Clin Radiol, 32, 571-5.
- Kuma K, Matsuzuka F, Kobayashi A (1992). Outcome of long standing solitary thyroid nodules. World J Surg, 16, 583-7.
- Meissner WA, Adler A (1958). Papillary carcinoma of the thyroid: a study of the pathology of two hundred twenty-six cases. AMA Arch Pathol, 66, 518-25.
- Moon WJ, Jung SL, Lee JH, et al (2008). Benign and malignant thyroid nodules: Ultrasound differentiation-multicenter retrospective study. Radiology, 247, 762-7.
- Nam-Goong IS, Kim HY, Gong G (2004). Ultrasonographyguided fine-needle aspiration of thyroid incidentaloma: correlation with pathological findings. Clin Endocrinol, 60, 21-8.
- Papini E, Guglielmi R, Bianchini A (2002). Risk of malignancy in nonpalpable thyroid nodules: predictive value of ultrasound and color doppler features. J Clin Endocrinol Metab, 87, 1941-6.
- Pusztaszeri MP, Sadow PM, Faquin WC (2013) Images in endocrine pathology: psammomatoid calcifications in oncocytic neoplasms of the thyroid, a potential pitfall for papillary carcinoma. Endocr Pathol, 13, 9242-2.
- Rahimi M, Farshchian N, Rezaee E, Shahebrahimi K, Madani H (2013). To differentiate benign from malignant thyroid nodule comparison of sonography with FNAC findings. Pak J Med Sci, 29, 77-80.
- Shin S, Park SE, Kim SY, et al (2014). Effectiveness of ultrasonographic screening for thyroid cancer: round-table conference in the national evidence-based healthcare collaborating agency (NECA) in conjunction with the Korean thyroid association. Asian Pac J Cancer Prev, 15, 5107-10.
- Solbiati L, Osti V, Cova L, Tonolini M (2001). Ultrasound of

- thyroid, parathyroid glands and neck lymph nodes. Eur Radiol, 11, 2411-24.
- Sushel C, Khanzada TW, Zulfikar IM, Samad A (2009). Histopathological pattern of diagnosis in patients undergoing thyroid operations. Rawal Med J, 34, 14-6.
- Taddesse A, Yaqub A (2011). Clinical, sonographic and cytological evaluation of small versus large thyroid nodules. JPMA, 61, 466-69.
- Unnikrishnan AG, Kalra S, Baruah M, Nair G, N Vasantha et al (2011). Endocrine society of india management guidelines for patients with thyroid nodules: a position statement. Indian J Endocrinol Metab, 15, 2-8.
- Vinayak S, Sande JA (2012). Avoiding unnecessary fine-needle aspiration cytology by accuractely predicting the benign nature of thyroid nodules using ultrasound. J Clin Imaging Sci, 2, 1-7.
- Wang N, Xu Y, Ge C, Guo R, Guo K (2006). Association of sonographically detected calcification with thyroid carcinoma. Head Neck, 28, 1077-83.
- Welker MJ, Orlov D (2003). Review thyroid nodules. Am Family Physician, 67, 559-66.
- Wienke JR, Chong WK, Fielding JR, Zou KH, Mittelstaedt CA (2003). Sonographic features of benign thyroid nodules: interobserver reliability and overlap with malignancy. J Ultrasound Med, 22, 1027-31.
- Wiest PW, Hartshorne MF, Inskip PD (1998). Thyroid palpation versus high-resolution thyroid ultrasonography in the detection of nodules. J Ultrasound Med, 17, 487-96.
- Yang Y, Li Q, Guo L, et al (2011). Retrospective analysis of thyroid cancer in China. Asian Pac J Cancer Prev, 12,
- Yunus M, Ahmed Z (2010). Significance of ultrasound features in predicting solid thyroid nodules: need for fine-needle aspiration. JPMA, 60, 848-53.
- Zuberi LM, Yawar A, Islam N, Jabbar A (2004). Clinical presentation of thyroid cancer patients in Pakistan-AKUH experience. JPMA, 54, 526-8.