Med Ultrason
2012, Vol. 14, no. 4, 271-277

The value of a new score for sonoelastographic differentiation between benign and malignant cervical lymph nodes

Lavinia Manuela Lenghel[1], Sorana D. Bolboacă[2], Carolina Botar-Jid[1], Grigore Băciuț[3], Sorin M. Dudea[1]

[1]Department of Radiology, Emergency Clinical County Hospital, [2]Department of Medical Informatics and Biostatsitics, [3]Department of Cranio-Maxillo-Facial Surgery "Iuliu Haţieganu" University of Medicine and Pharmacy, Cluj-Napoca

Abstract

Objective: The aim of this study is to explore the diagnostic value of sonoelastography for the differentiation between benign and malignant superficial lymph nodes of the neck. In this respect the utility of an original scoring system was explored. **Material and method:** Over a period of 30 months the patients examined routinely for the assessment of superficial lymph nodes of the neck were recorded in a data base containing grey-scale, Doppler and sonoelastographic information and images. The sonoelastographic images of 30 benign and 39 malignant lymph nodes were assessed. The images were scored according to a new, eight pattern scoring system proposed by our group. Interobserver agreement and area under the ROC curve (AUROC) for the differentiation between benign vs. malignant and benign vs. metastatic nodes were analyzed. **Results:** The analysis of the interobserver agreement for the investigated score provided a weighted Kappa = 0.687, 95%CI [0.572 to 0.802] and standard error = 0.059. In the differentiation benign – malignant, the AUROC was 0.846, with sensitivity of 66.67% and specificity of 96.67% for score >3. In the differentiation between benign and metastasis, the same criterion provided an AUROC of 0.855, with sensitivity of 71.43 and specificity of 96.67%. **Conclusions:** Our study suggests that applying the proposed score provides good interobserver agreement. The score also provided very good specificity and reasonable sensitivity in the differentiation between malignant and benign lymph nodes in the neck.

Keywords: ultrasonography, elastography, lymph node, neck

Introduction

Ultrasonographic evaluation of the cervical lymph nodes always follows the clinical exam and is very important for the diagnosis, prognosis and selection for the appropriate treatment [1]. The head and neck contains 60 to 70 lymph nodes of the 400 to 450 in the human body. The detection and description of the characteristics of the lymph nodes are important due to a large number of inflammatory or neoplastic diseases, primary or secondary, that can involve cervical lymph nodes [2].

Frequently, metastatic lymph nodes are the first sign in patients with malignancies. The most important task for imaging techniques is to differentiate between malignant and inflammatory lymphadenopathy. Ultrasonography is the first imaging technique used for the differentiation between benign and malignant cervical lymph nodes [1]. A study by Xing et al [3] proved that ultrasound is the most sensitive and specific examination, with the highest predictive value in detection and diagnosis of melanoma metastatic cervical lymphadenopathy.

Routinely, grey-scale and Doppler images are used for the diagnosis of lymph nodes. The main grey-scale criteria for benign-malignant differentiation are: size, shape, margins, echogenicity, presence of hilum, necrosis, calcifications, matting, and surrounding edema. Doppler criteria include presence and distribution of the flow,

Received 12.09.2012 Accepted 18.10.2012
Med Ultrason
2012, Vol. 14, No 4, 271-277
Corresponding author: Lavinia Manuela Lenghel
Radiology Department,
"Iuliu Haţieganu" University of Medicine and Pharmacy
1-3 Clinicilor str, 400006
Cluj-Napoca, Romania
E-mail: manu_2416@yahoo.com

vascular pattern, number of vascular pedicles and impedance values (resistence index - RI, pulsatility index - PI). However, as far as diagnostic accuracy is concerned, either sensitivity or specificity was not satisfactory for any of the above mentioned signs [4-12].

Recently, sonoelastography was introduced in the clinical practice as a diagnostic tool used for the assessment of the diseases of the organs amenable to a small distance from the transducer.

The purpose of our study was to explore the diagnostic value of sonoelastography for the differentiation between benign and malignant superficial lymph nodes of the neck. In this respect the utility of an original scoring system was explored. To achieve the purpose we targeted the following goals:

– to assess the interobserver variability of the score;

– to determine the sensitivity and specificity of differentiating between benign and malignant nodes based on the consensus scoring attributed to the nodes aiming to determine which cutoff level provides the highest sensitivity and/or specificity. For this purpose the area under the receiver-operator curve (AUROC) for different cutoff levels for our scoring systems was analyzed;

– to assess the diagnostic value of the score only for the differential diagnosis between benign and metastatic nodes, eliminating lymphomas.

Material and methods

Over a period of 30 months (October 2009 – April 2012) the patients examined routinely for the assessment of superficial lymph nodes in the Division of Ultrasound of the Department of Radiology were recorded in a database containing grey-scale, Doppler and sonoelastographic information (such as demographics, size, echogenicity, margins, structure, homogeneity and internal structure changes, Doppler signal presence and distribution, RI, PI and node stiffness) and images.

Follow-up information or pathology reports were also recorded, whenever available. Care was taken to match the pathology reports of individual nodes with the ultrasonographic images. All the images were acquired by one examiner (ML).

The study was retrospective and randomized. The randomization consisted of the anonymization and numbering of images, by a researcher unaware of the clinical outcome. The number code was used to split the images with known outcome into two categories: benign and malignant, by personnel aware of the coding but not involved in the study. The study was approved by the ethical committee of the institution. Due to the retrospective design of the study written consensus was waived.

One member of the team randomly chose the images, by numbers, the only available information being about the final outcome (benign or malignant).

The sonoelastographic images of 30 benign and 39 malignant lymph nodes were chosen. The inclusion criteria were:

– good quality representative elastographic images.

– follow-up of the least 6 months with no change or regression of the lymph node for benign lymph nodes.

– availability of pathology for the malignant nodes.

– for the malignant nodes, matching between the imaged lymph node and the node reported at pathology.

Patients without follow-up, no pathological report, non-matching criteria between the image and pathology and technically inadequate sonoelastographic images such as inappropriate compression quality factor, incomplete encompassing of the node or insufficient perinodal tissue for comparison were excluded.

The images were obtained on a Hitachi EUB 8500 scanner with 6.5-13 MHz linear transducers and with real time sonoelastography. The nodes were scanned on longitudinal and transverse planes with wide color window, trying to encompass as much soft tissue around the lymph node as possible. Probe stabilizer was not used. Compression was applied perpendicular to the skin. Only the images with an acquisition quality factor of 3 or 4, as displayed by the specific scale of the scanner, were stored.

The images were scored according to a scoring system proposed by our group, as follows:

- pattern 1: the whole nodule is soft (similar to the surrounding tissues);
- pattern 2: <50% of the nodule surface is blue and no individualized hypoechoic nodules are seen in the lymph node structure;
- pattern 3: <50% of the nodule surface is blue and individualized soft hypoechoic nodules are seen in the lymph node structure;
- pattern 4: <50% of the nodule surface is blue and individualized hard hypoechoic nodules are seen in the lymph node structure;
- pattern 5: 50% - 100% of the nodule surface is blue and no individualized hypoechoic nodules are seen in the lymph node structure;
- pattern 6: 50% - 100% of the nodule surface is blue and individualized hard hypoechoic nodules are seen in the lymph node structure;
- pattern 7: the blue area covers the whole nodules and extends in the soft tissues
- pattern 8: blue (hard) nodule containing fluid areas (necrosis).

The scoring system is illustrated in figure 1.

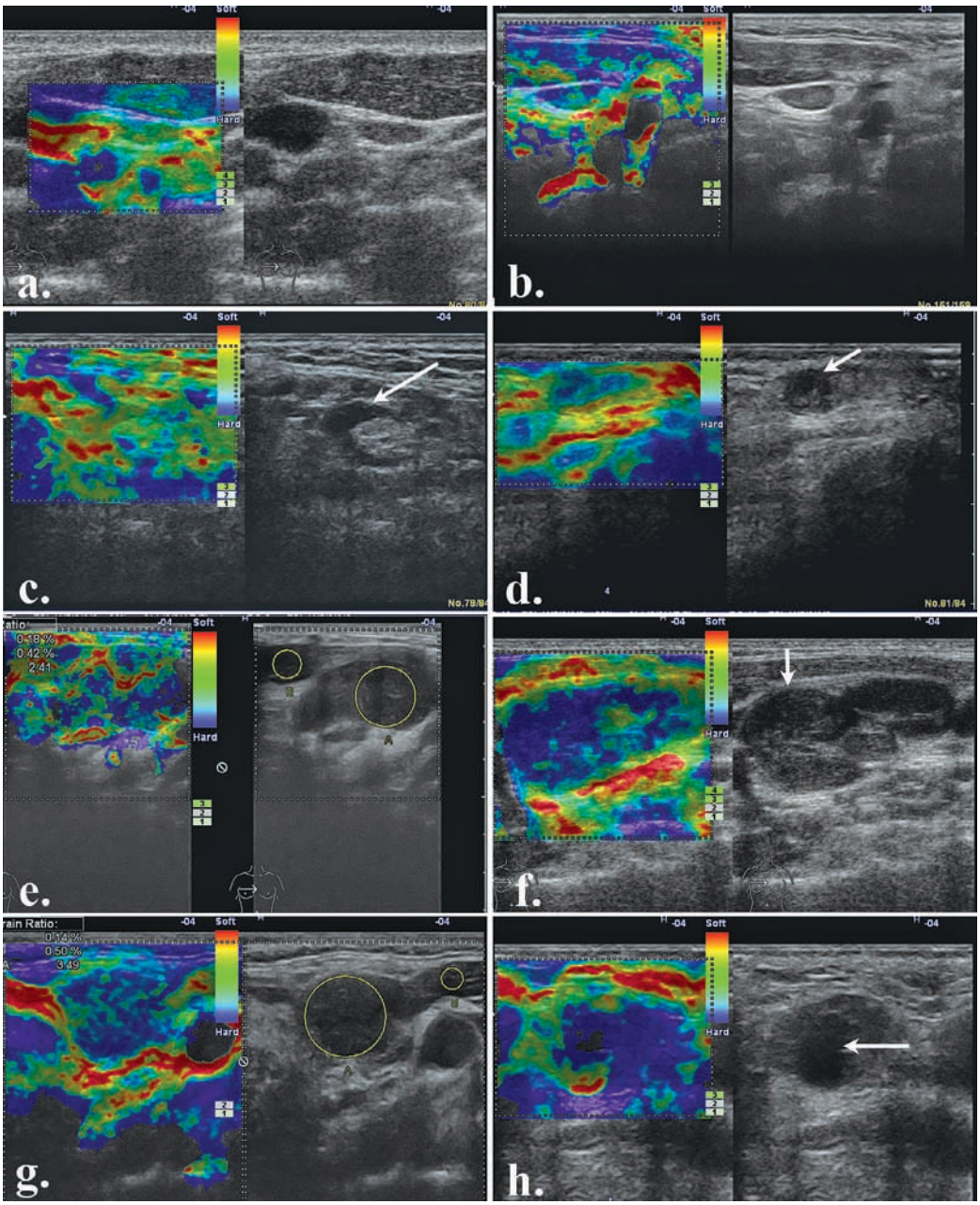


Fig 1. Illustration of the scoring system: a) pattern 1 – the whole nodule is soft, similar to the surrounding tissues; b) pattern 2 – <50% of the nodule surface is blue and no individualized hypoechoic nodules are seen in the lymph node structure; c) pattern 3 – <50% of the nodule surface is blue and individualized soft hypoechoic nodules are seen in the lymph node structure (arrow); d) pattern 4 – <50% of the nodule surface is blue and individualized hard hypoechoic nodules are seen in the lymph node structure (arrow); e) pattern 5 – 50% – 100% of the nodule surface is blue and no individualized hypoechoic nodules are seen in the lymph node structure; f) pattern 6 – 50% – 100% of the nodule surface is blue and individualized hypoechoic hard nodules are seen in the lymph node structure (arrow); g) pattern 7 – the blue area covers the whole nodule and extends in the soft tissues; h) pattern 8 – blue (hard) nodule containing fluid areas (necrosis) (arrow).

The time interval between the acquisition of the last image and image analysis was at least 3 months. Two examiners (ML and SD) with similar experience in sonoelastography (4 respectively 6 years of practice) assessed independently the images applying the scoring system for each image under investigation. A consensus session of the two examiners led to the final score attributed to each image.

Statistical methods. The interobserver agreement was assessed using k weighted statistics [13]. The interpretation of the strength of agreement is: poor if $k < 0.20$; fair if $0.21 < k < 0.40$, moderate if $0.41 < k < 0.60$, good if $0.61 < k < 0.80$, and very good if $0.81 < k < 1$ [14]

The diagnostic value (sensitivity and specificity) for each score were assessed by exploring the AUROC [15].

The diagnostic value was assessed first for the benign – malignant differential diagnosis. A second analysis for the differential between benign and metastasis was performed after removing the lymphoma nodes.

Results

The 69 lymph nodes were recorded in 42 patients. In the malignant category, there were 11 lymphomatous and 28 metastatic lymph nodes.

The demographic characteristics of the studied lot are presented in table I.

The analysis of the interobserver agreement for the investigated score provided a weighted Kappa = 0.687, 95%CI [0.572 to 0.802] and standard error = 0.059.

The pattern distribution of the 69 lymph nodes, after the consensus evaluation, is presented in table II.

The AUROC analysis in the differentiation between benign and malignant nodes is presented in table III and illustrated in figure 2. The nodes with lymphoma and metastases (n=39) were considered positive. The nodes with inflammatory changes (n=30) were considered negative. The AUROC was 0.846, standard error (SE) =0.0437 with a 95% confidence interval (CI) 0.739 – 0.921 and a significance level P (area =0.5) <0.0001.

Table I. Demographics of the studied group

Diagnosis	N patients	N lymph nodes	Male		Age - years (mean, extremes)	Female		Age - years (mean, extremes)
			N patients	N nodes		N patients	N nodes	
Benign	22	30	9	12	26.77 (22-77)	13	18	31 (20-70)
Lymphoma	6	11	3	7	34 (22-54)	3	4	26 (21-33)
Malignant	14	28	10	12	26.15 (48-70)	4	16	27.37 (32-70)

N- number

Table II. Pattern distribution of the nodes

Pattern	1	2	3	4	5	6	7	8
N nodes	13	21	8	3	11	5	6	2

N- number

Table III. AUROC analysis of the investigated score for the diagnosis benign - malignant

Criterion	Sensitivity	95% CI	Specificity	95% CI	+LR	-LR
>=1	100.00	91.0 - 100.0	0.00	0.0 - 11.6	1.00	
>1	97.44	86.5 - 99.9	40.00	22.7 - 59.4	1.62	0.064
>2	69.23	52.4 - 83.0	73.33	54.1 - 87.7	2.60	0.42
>3 *	66.67	49.8 - 80.9	96.67	82.8 - 99.9	20.00	0.34
>5	56.41	39.6 - 72.2	96.67	82.8 - 99.9	16.92	0.45
>6	30.77	17.0 - 47.6	100.00	88.4 - 100.0		0.69
>8	0.00	0.0 - 9.0	100.00	88.4 - 100.0		1.00

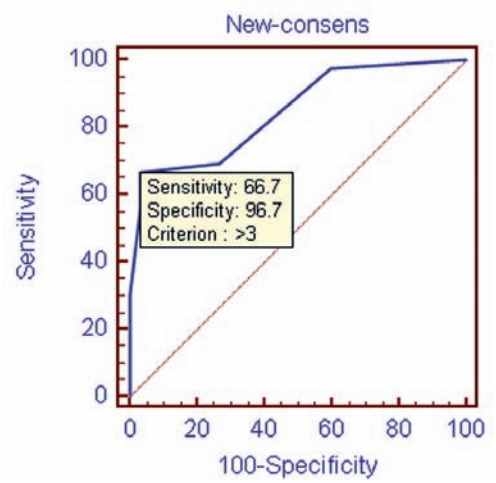


Fig 2. The AUROC analysis in the differentiation between benign and malignant nodes

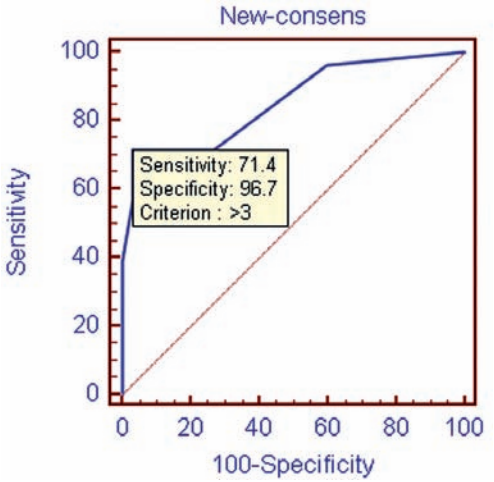


Fig 3. The AUROC analysis in the differentiation between benign and metastatic nodes

Table IV. AUROC analysis of the investigated score for the diagnosis benign - metastasis

Criterion	Sensitivity	95% CI	Specificity	95% CI	+LR	-LR
>=1	100.00	87.7 - 100.0	0.00	0.0 - 11.6	1.00	
>1	96.43	81.7 - 99.9	40.00	22.7 - 59.4	1.61	0.089
>2	71.43	51.3 - 86.8	73.33	54.1 - 87.7	2.68	0.39
>3 *	71.43	51.3 - 86.8	96.67	82.8 - 99.9	21.43	0.30
>5	57.14	37.2 - 75.5	96.67	82.8 - 99.9	17.14	0.44
>6	39.29	21.5 - 59.4	100.00	88.4 - 100.0		0.61
>8	0.00	0.0 - 12.3	100.00	88.4 - 100.0		1.00

The AUROC analysis in the differentiation between benign and metastatic nodes is presented in table IV and illustrated in figure 3. The nodes with metastases (n=28) were considered positive and the nodes with inflammatory changes (n=30) were considered negative. The AUROC was 0.855, SE = 0.0486 with a 95% CI = 0.738 – 0.934 and a significance level P (area =0.5) <0.0001.

Discussion

Elastography is a new dynamic, ultrasonographic technique, recently developed, that has the possibility of providing an estimation of tissue stiffness and, therefore, has great potential to become the necessary tool for differentiating benign from malignant lymph nodes.

The results of our study showed good agreement between the two observers regarding the lymph nodes belonging to a specific pattern. From this point of view the result is encouraging as the application of this score did not show a high observer dependency.

The cutoff line between benign and malignant lymph nodes in the score developed by our group is situated between pattern 3 and pattern 4, meaning that a hypoechoic hard nodule into a lymph node is more likely to indicate tumor infiltration, even if the rest of the nodule has an inflammatory aspect. This score reached high specificity, but the sensitivity is lower than the one reported for other scores. This could be due to the inclusion of a less suggestive aspect for malignancy, but it can also be a consequence of the quantitative and/or qualitative criteria used for the assessment of the images. Specificity remains high and sensitivity slightly increases when lymphomas were removed from the studied group. Because of the large size of the lymph nodes, their homogeneous structure and low stiffness, from the elastographic point of view, lymphomatous nodes are more likely to show a benign aspect.

Lyshchik et al [16] studied the stiffness of cervical lymph nodes using sonographic grey-scale elastography. Strain index >1.5 proved to be the most useful sign for detecting metastatic lymph nodes (98% specificity, 85% sensitivity and 92% accuracy).

A paper published in 2007 by Furukawa et al [17] proposed an elasticity score of four patterns applied to metastatic lymph nodes from head and neck squamous cell carcinoma. The score contains the following patterns: pattern 1- 80% or more of the cross-sectional are of the lymph node is red or green, i.e. soft; pattern 2- 50% or more an less than 80% is red or green; pattern 3- 50% or more an less than 80% is blue; pattern 4- 80% or more of the cross-sectional are of the lymph node is blue, i.e. hard. However sensitivity and specificity as well as a cut-off pattern between benign and malignant lymph nodes were not defined.

In 2008, Alam et al [18] evaluated the accuracy of B-mode sonography and sonographic elastography individually and then combined, for the differentiation between benign and malignant lymph nodes. The elasticity score developed by Alam et al consists of five patterns: pattern 1 - absent or very small blue area(s); pattern 2 - small scattered blue areas, total blue area < 45%; pattern 3 - large blue area(s), total blue area >/= 45%; pattern 4 - peripheral blue area and central green area, suggesting central necrosis; pattern 5- blue area with or without a green rim. The cutoff line benign-malignant was set between score 2 and 3. Elastography showed a sensitivity of 83%, a specificity of 100% and an accuracy of 89% (using combined quantitative and qualitative criteria).

Another elasticity score used for the assessment of lymph nodes is the score proposed by Rago et al [19]. Initially, this score was proposed for thyroid nodules, but it was applied also in enlarged cervical lymph nodes [20]. The elasticity score proposed is: score 1 - elasticity in the whole nodule; score 2 - elasticity in a large part of the nodule; score 3 - elasticity only at the peripheral part of the nodule; score 4 - no elasticity in the nodule; score 5 - no elasticity in the nodule and in the posterior shadowing. The elasticity scores 4 and 5 were highly predictive for malignancy (97% sensitivity, 100% specificity, predictive positive value 100%, negative predictive value 98%).

The scores mentioned above imply some issues about their clinical applicability. The score proposed by Furukawa et al [17] does not estimate a diagnostic value for sensitivity and specificity. The scores developed by Alam et al [18] and Rago et al [19] assess only the sonoelastographic aspect of the nodule, without a direct correlation with the grey-scale appearance. The combination between the sonoelastographic image and the presence of a hypoechoic nodule is described and used in prostate elastography [21]. None of the scores mentioned above includes a structural change - focal cortical hypertrophy. This particular finding can indicate tumor infiltration in an initially benign lymph node, but it can be also observed in acute reactive lymphadenopathy. The score proposed by our group tried to include these findings, by elastographic characterization of the focal cortical nodules and comparing the hypoechoic nodules seen in grey-scale with the elastographic findings to determine the nature of the nodules.

The lymph node surface percentage affected by stiffness (45% or 80%) is difficult to assess by simple inspection of the image. The reason why our group proposed 50% of the surface is that it is much easier for the observer to inspect the image and include the lymph node appearance in the correct pattern. Apart from the score proposed by Alam et al [18], which includes the suggestion of necrosis, none of the other elastographic scores mention the intranodal structural changes. The new score proposed by our group includes the intranodal liquefaction changes seen on 2D images observed in metastatic nodules, highly suggestive for malignancy.

Considering the results, it would be interesting to compare the diagnostic performance of the score proposed by our group with the scores mentioned in the literature, on the same group of patients.

The limitations of the new score proposed by our group arise from the small number of studied cases (the lowest number required for statistical significance) and, also, from the fact that the lymph node, not the patient, is used as the study unit. To the best of our knowledge, there are no studies that investigate the dependence of the elastographic appearance to the size of the nodule. Another possible limitation could be represented by the bias induced by the fact that the images were obtained by the same researcher. To reduce the effects of this bias, the images were anonymized and the time interval between the acquisition of the last image and image analysis was at least 3 months. The study did not try to differentiate directly between benign and malignant lymph nodes, an aspect that could have been biased by the acquisition of the images; it aimed at the assignment of a score to each individual image, irrespective of the final – and unknown – outcome. Therefore, the possible preexisting information had no relevance. Because of the reduced number of cases there were a small number of lymph nodes included in certain scores.

Conclusion

Our study suggests that applying the proposed score provides a good interobserver agreement. The score also

provided very good specificity and reasonable sensitivity in the differentiation between malignant and benign lymph nodes in the neck while taking into account not only the pure elastographic images but also the grey-scale appearance.

Further studies are required in order to compare the diagnostic value of the existing elastographic scores and to improve the differential diagnosis of abnormal lymph nodes.

Conflict of interest: none

References

1. Badea RI, Băciuț M. Limfonodulii cervicali. In: Badea RI, Dudea SM, Mircea PA, Zdrenghea D. Tratat de ultrasonografie clinică. vol 2. Capul şi gâtul, toracele şi mediastinul, ecocardiografie şi vasele membrelor. Ed. Medicală, Bucureşti 2006:133-147.
2. Das D, Gupta M, Kaur K, Kalucha A. Elastography: the next step. J Oral Sci 2011;53:137-141.
3. Xing Y, Cromwell KD, Cormier JN. Review of diagnosis imaging modalities for the surveillance of melanoma patients. Dermatol Res Pract 2012;2012:941-921.
4. Ahuja AT, Ying M. Sonographic evaluation of cervical lymph nodes. AJR Am J Roentgenol 2005;184:1691-1699.
5. Ahuja A, Ying M, King A, Yuen HY. Lymph node hilus: gray scale and power Doppler sonography of cervical nodes. J Ultrasound Med 2001;20:987-92.
6. Ying M, Ahuja A, Brook F, Metreweli C. Vascularity and grey-scale sonographic features of normal cervical lymph nodes: variations with nodal size. Clin Radiol 2001;56:416-419.
7. Ahuja A, Ying M. Sonography of neck lymph nodes. Part II: abnormal lymph nodes. Clin Radiol 2003;58:359-366.
8. Ahuja AT, Ying M, Ho SS, Metreweli C. Distribution of intranodal vessels in differentiating benign from metastatic neck nodes. Clin Radiol 2001;56:197-201.
9. Ahuja A, Ying M. Sonographic evaluation of cervical lymphadenopathy: is power Doppler sonography routinely indicated? Ultrasound Med Biol 2003;29: 353-359.
10. van den Brekel MW, Castelijns JA, Snow GB. The size of lymph nodes in the neck on sonograms as a radiologic criterion for metastasis: how reliable is it? AJNR Am J Neuroradiol 1998;19:695-700.
11. Khanna R, Sharma AD, Khanna S, Kumar M, Shukla RC. Usefulness of ultrasonography for the evaluation of cervical lymphadenopathy. World J Surg Oncol 2011;9:29.
12. Ahuja AT, Ying M, Ho SY, et al. Ultrasound of malignant cervical lymph nodes. Cancer Imaging 2008;8:48-56.
13. Fleiss JL, Levin B, Paik MC. Statistical methods for rates and proportions. 3rd ed. John Wiley & Sons, Hoboken 2003.
14. Altman DG. Practical statistics for medical research. Chapman and Hall, London 1991.
15. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 1988;44:837-845.
16. Lyshchik A, Higashi T, Asato R, et al. Cervical lymph node metastases: diagnosis at sonoelastography – initial experience. Radiology 2007;243:258-267.
17. Furukawa MK, Fujita Y, Kubota A, Furukawa M, Hanamura H. Diagnosis of cervical lymph node metastasis of head and neck squamous cell carcinoma - usefulness of Power Doppler ultrasonography and elastography. Medix 2007;Suppl:20-23.
18. Alam F, Naito K, Horiguchi J, Fukuda H, Tachikake T, Ito K. Accuracy of sonographic elastography in the differential diagnosis of enlarged cervical lymph nodes: comparison with conventional B-mode sonography. AJR Am J Roentgenol 2008;191:604-610.
19. Rago T, Santini F, Scutari M, Pinchera A, Vitti P. Elastography: new developments in ultrasound for predicting malignancy in thyroid nodules. J Clin Endocrinol Metab 2007;92:2917-2922.
20. Moon HJ, Kim EK, Yoon JH, Kwak JY. Clinical implication of elastography as a prognostic factor of papillary thyroid microcarcinoma. Ann Surg Oncol 2012;19:2279-2287.
21. Dudea SM, Giurgiu CR, Dumitriu D, et al. Value of ultrasound elastography in the diagnosis and management of prostate carcinoma. Med Ultrason 2011;13:45-53.