ORIGINAL ARTICLE

Comparison of Two Elasticity Scoring Systems in the Assessment of the Cervical Lymph Nodes

CME Credits   CrossMark

Wu-Chia Lo [1,2,3*], Li-Jen Liao [1,2,4]

[1] Department of Otolaryngology, Far Eastern Memorial Hospital, [2] Department of Healthcare Administration, Oriental Institute of Technology, [3] Department and Graduate Institute of Pathology, National Taiwan University College of Medicine, and [4] Graduate Institute of Epidemiology and Preventive Medicine, College of Public Health, National Taiwan University, Taipei, Taiwan

Background: Real-time elastography (RTE) and elasticity scoring system is a new diagnostic tool in the evaluation of enlarged cervical lymph nodes (LNs). The aim of the study was to investigate the validity and reliability of two elasticity scoring systems in discriminating cervical lymphadenopathies.
Materials and methods: The study protocol was approved by the institutional review board of the local ethics committee. Ninety-eight consecutive patients underwent ultrasound-guided fine needle aspiration biopsy after conventional ultrasound and RTE evaluation. Elasticity scores were divided into four and five categories according to signal distribution, based on previously determined models. Two independent raters analyzed the elastographic data.
Results: Final histology revealed 60 benign and 38 malignant cervical LNs. In the four-point scoring system, 41/60 benign LNs were classified as score 1 or 2 and 28/38 malignant nodes as score 3 or 4 [sensitivity 73.7%, specificity 68.3%, positive predictive value (PPV) 59.6%, negative predictive value (NPV) 80.4%]. Using the five-point scoring system, 28/60 benign nodes were classified as score 1 or 2 and 31/38 malignant nodes as score 3—5 (sensitivity 79.0%, specificity 45.0%, PPV 47.6%, NPV 80.4%). In the four-point scoring system, rates of concordance between raters ($\kappa = 0.63$) and with the same rater ($\kappa = 0.75$) were both good with statistical significance (both $p < 0.01$). With the five-point scoring system, both kappa statistic tests between raters ($\kappa = 0.68$) and with the same rater ($\kappa = 0.68$) were also good with statistical significance (both $p < 0.01$).
Conclusion: For qualitative RTE, the four-point scoring system had value similar to the five-point scoring system in predicting malignancy in cervical LNs. Furthermore, the reliability

Conflicts of interest: None.
* Correspondence to: Dr Wu-Chia Lo, Department of Otolaryngology, Far Eastern Memorial Hospital, 21, Section 2, Nan-Ya South Road Pan Chiao, Taipei 220, Taiwan.
E-mail address: lowuchia@gmail.com or
E-mail address: d02444001@ntu.edu.tw (W.-C. Lo).

was comparable in both scoring systems. For the purpose of simplified evaluation, we suggested using the four-point scoring system to rate the qualitative RTE in the future.

Introduction

Palpable neck lymph nodes (LNs) can be categorized as benign or malignant enlargement. For the subsequent diagnosis and treatment, it is important to confirm that the LNs are malignant or not. High-resolution ultrasound (US) is a reliable and widely used imaging modality to assess an enlarged cervical LN [1,2].

Real-time elastography (RTE) is a novel technique to assess tissue elasticity (stiffness) by comparing local tissue displacements from ultrasonic signals before and after the application of a compressive force [3—10]. Soft tissues demonstrate more deformation than stiff tissues while compressing the transducer. Revealing that malignant tissues are stiffer than their surrounding benign counterparts, some recent reports document the feasibility of real-time US elastography to distinguish malignant from benign neck LNs [3—10]. The reported sensitivity, specificity, and accuracy of US elastography in the assessment of cervical LNs ranged from 50% to 88%, 35% to 100%, and 66% to 92%, respectively [3—10]. As a result, some authors believe US elastography is a promising imaging technique in differentiating benign and malignant neck LNs; however, others argue that it is incorrect to guide decision-making concerning selection of a LN for needle biopsy [3—10]. In addition, some researchers [3,5,6,9] use the four-point elasticity score (ES) and others [4,7] use the five-point ES to interpret the elastrographic data of enlarged neck LNs.

The current study aimed to evaluate the validity and reliability of real-time qualitative elastography, compared with four- and five-point scoring systems, in the differentiation of malignant and benign cervical LNs.

Patients and methods

The study was conducted at the Far Eastern Memorial Hospital, Taipei, Taiwan from May 2011 to September 2011. The study protocol was approved by the institutional review board of the local ethics committee (FEMH-IRB-100162-E). Each patient gave written informed consent. The inclusion criteria were palpable neck LNs and at least one sonographic characteristic of abnormality [2,11—14] (i.e., size, shape, margin, hilar echogenicity, or vascular pattern) that permitted US-guided fine needle aspiration biopsy (FNAB). The exclusion criteria were previous history of malignancy or enlarged cervical LNs related to known diseases, such as tuberculous lymphadenitis or Kikuchi disease. Eventually, sonograms of 98 cases were enrolled and analyzed. The sonograms were performed by two experienced sonologists (raters A and B) using a high-resolution 7—18-MHz real-time linear-array transducer (Aplio MX, Toshiba, Tokyo, Japan). All patients included in our study underwent thorough neck examination using gray-scale and power Doppler US. The lengths of short axes (S) and long axes (L), and diameter ratio of short to long axis (S/L ratio) of enlarged LNs were measured. The nodal margin was defined as regular or irregular. Echogenicity with respect to the adjacent muscles was evaluated and classified as hypoechoic, isoechoic, or hyperechoic. The presence or absence of echogenic hilum was determined. The internal echo was examined for the presence of homogenous or heterogeneous pattern. Vascular pattern was studied by power Doppler US and classified as the group of avascular or hilar type versus that of mixed, spotted, or peripheral type. Morphological US parameters were recorded on the Marosis PACS system (Marotech Inc., Seoul, South Korea).

All patients were also examined by real-time US elastography for an additional 2—3 minutes immediately after conventional US. For the elastographic technique, RTE images were displayed along with gray-scale sonograms as a two-panel image. Next, compression with light pressure followed by decompression was repeated until almost the same size and color distribution of the region of interest (ROI) in numerous sequential images was achieved. The direction of the compression was along the radiation axis while attention was focused on avoidance of out-of-plane motions. RTE images were a color-coded graphic representation of the selected target area, such that blue indicated stiff, green and yellow indicated intermediate stiffness, and red indicated soft. The elastographic pictures of target LNs were saved as bitmap files on a hard disk by the research nurse. At the conclusion of the examination, each patient underwent US-guided FNAB, with placement of the array probe parallel to the needle for the guidance and position of a 22-gauge needle within the node that was suspected of malignancy. Final diagnosis was based on the basis of cytopathologic and histopathologic findings. Patients who had negative cytopathologic diagnoses were followed for a minimum of 6 months to verify that no subsequent malignancy had developed in these LNs.

The elastograms were reviewed independently by two raters (A and B) who had performed elastography. Elastographic images were assessed anomynously without any clinical data after randomization. The four- and five-point elasticity scores (ES) used in the current study to assess the ratio of stiff parts within the LN were adapted from Bhatia et al [5] (Table 1) and Alam et al [4] (Table 2), respectively. Rater A evaluated the entire elastographic image twice with an interval of 1 month, and rater B assessed the data once.

Statistical analysis

Mann-Whitney *U* tests and Chi-square tests were used to determine the differences in clinical parameters (i.e., age,