Real-time prediction of mediastinal lymph node malignancy by endobronchial ultrasound

Hanaa Shafiek, a,c Federico Fiorentino, a Alejandro David Peralta, a Enrique Serra, b Blanca Esteban, b Rocío Martínez, d María Angles Noguera, a Pere Moyano, a Ernest Sala, a,d Jaume Sauleda, a,d Borja G. Cosío a,d,*

a Departamento de Medicina Respiratoria, Hospital Universitario Son Espases, Palma de Mallorca, Spain b Departamento de Anatomía Patológica, Hospital Universitario Son Espases, Palma de Mallorca, Spain c Departamento de Enfermedades del Tórax, Facultad de Medicina, Universidad de Alejandroía, Alejandroía, Egypt d Ciber de Enfermedades Respiratorias (Ciberes), Spain e Departamento de Anestesiología, Hospital Universitario Son Espases, Palma de Mallorca, Spain

ABSTRACT

Objective: To evaluate the utility of different ultrasonographic (US) features in differentiating benign and malignant lymph node (LN) by endobronchial ultrasound (EBUS) and validate a score for real-time clinical application.

Methods: A total of 208 mediastinal LNs acquired from 141 patients were analyzed. Six different US criteria were evaluated (short axis ≥ 10 mm, shape, margin, echogenicity, central hilar structure [CHS], and presence of hyperechoic density) by two observers independently. A simplified score was generated where the presence of margin distinction, round shape and short axis ≥ 10 mm were scored as 1, and heterogeneous echogenicity and absence of CHS were scored as 1.5. The score was evaluated prospectively for real-time clinical application in 65 LNs during EBUS procedure in 39 patients undertaken by two experienced operators. These criteria were correlated with the histopathological results and the sensitivity, specificity, and positive and negative predictive values (PPV and NPV) were calculated.

Results: Both heterogeneity and absence of CHS had the highest sensitivity and NPV (> 90%) for predicting LN malignancy with acceptable inter-observer agreement (92% and 87% respectively). On real-time application, the sensitivity and specificity of the score ≥ 5 were 78% and 86% respectively; only the absence of CHS, round shape and size of LN were significantly associated with malignant LN.

Conclusions: A combination of different US criteria can be useful for the prediction of mediastinal LN malignancy and valid for real-time clinical application.

© 2014 SEPAR. Published by Elsevier España, S.L. All rights reserved.

Predicción en tiempo real de la malignidad de ganglios linfáticos mediastínicos mediante ecografía endobronquial

RESUMEN

Objetivo: Evaluar la utilidad de diferentes características ecográficas para diferenciar los ganglios linfáticos (GL) benignos y malignos mediante ecografía endobronquial (EBUS) y validar una puntuación para una aplicación clínica en tiempo real.

Métodos: Se analizaron 208 GL mediastínicos procedentes de 141 pacientes. Dos observadores evaluaron de manera independiente 6 criterios ecográficos diferentes (eje menor ≥ 10 mm, forma, margen, ecogenicidad y estructura hilar central [EHC] y presencia de densidad hiperecogénica). Se generó una puntuación simplificada en la que a la presencia de márgenes bien definidos, la forma redondeada y el eje menor

* Corresponding author.
E-mail address: borja.cosio@ssib.es (B.G. Cosío).
Introduction

Lymph node (LN) biopsy is essential for accurate staging of lung cancer. Endobronchial ultrasound (EBUS) is a technique that combines endoscopic visualization with high frequency ultrasound imaging. EBUS is useful for visualizing both mediastinal and hilar LNs and for guiding needle aspiration during the cytological and histological biopsy procedure.

In the last decade, many studies have evaluated sonographic observations that may suggest malignant LN infiltration in head and neck, breast, uterine cervix and esophageal cancer. In a recent retrospective study in lung cancer, Fujiwara et al. described 4 sonographic features of value for predicting the malignant infiltration of mediastinal LN (round shape, distinct margin, heterogeneous echogenicity and the presence of coagulation necrosis sign). Schmid-Bindert et al. proposed a score that combined the following ultrasound criteria for predicting malignancy in mediastinal LN: round shape, well-defined margins, echogenicity, the absence of any central hilar structure (CHS), short axis ≥ 1 cm and color power Doppler index grade 2 or 3. The authors concluded that if less than 3 of the specified criteria were present, the LN had a very low probability of being malignant. However, these criteria are retrospective and may have a subjective component, since no prospective validation was made and some of the criteria used are difficult for experienced operators to determine in real time.

In an attempt to validate the clinical utility of the foregoing criteria, the hypothesis was proposed that some LN ultrasound criteria could be useful for predicting malignant LN infiltration and thus could be used as a guide for selecting LNs for aspiration, leading to improved bronchogenic staging. Consequently, the objectives of this study were, firstly, to evaluate retrospectively which EBUS sonographic features were the most accurate for differentiating benign involvement from malignant involvement in both mediastinal and hilar LNs, after correlating these features with the histopathological results; and secondly, to prospectively analyze the previously proposed score for a combination of these ultrasound criteria for real-time clinical application.

Methods

Design and study patients

The study was divided into 2 parts. The first was a descriptive study for which 176 patients with mediastinal/hilar LN undergoing EBUS in the Hospital Universitario Son Espases between 2009 and 2012 for lung cancer staging or investigation of suspected malignant LN infiltration were screened. Of these patients, 141 were included in the study, and 2 raters who were unaware of the final diagnosis independently analyzed the 208 LN images. The second part was a prospective study for which 39 consecutive patients were recruited and two raters predicted the probability of malignant LN involvement in real time on the basis of the previously validated criteria. Consecutive patients scheduled to undergo EBUS to study mediastinal/hilar LN observed on a computed tomography (CT) of the chest or hypermetabolic LN identified in a positron emission tomography with fluorodeoxyglucose (PET-FDG) were included.

The indication for performing invasive mediastinum staging by EBUS, endoscopic ultrasound (EUS), or a combination of both techniques is established by a local multidisciplinary lung cancer committee, according to the following standardized criteria:

(a) enlarged discrete mediastinal LN with PET uptake; (b) PET activity in a mediastinal LN and nodes with normal appearance on CT, and (c) suspected N2,3 involvement with a radiographically normal mediastinum (detected by CT and PET) and a central or N1 tumor. Patients with any contraindication for bronchoscopy according to the recommendations of the American Thoracic Society or a high anesthetic risk (American Society of Anesthesiologists physical status ≥ 3) were excluded. A full clinical history was obtained and chest CT and PET-FDG images were analyzed by a multidisciplinary lung cancer committee that established the indication for performing EBUS with transbronchial needle aspiration (TBNA) or transtracheal aspiration (TTA), following international guidelines. Patients were then evaluated by the anesthesiology team before the intervention. Informed consent was obtained from all patients.

EBUS + TTA/TBNA procedure

EBUS was performed using an EB-1970UK (Pentax, 10.0-5.0MHz, Tokyo, Japan) endoscope and the Hitachi Digital Ultrasound Scanner EUB-7000HV. The procedure was performed via the oral route under deep sedation with propofol, midazolam and fentanyl with anesthesiologist-controlled spontaneous ventilation in all study patients. The intervention was performed as described by Yasufuku et al. Enlarged LNs (greater than 5 mm) were identified by the measurements made in the digitally captured ultrasound images. Fixed images were captured for subsequent analysis. Blood vessels were confirmed with the Doppler modality. For TBNA and/or TTA, a 22-caliber needle was used. The needle was introduced via the working channel of the EBUS through the bronchial wall, and the LN was aspirated under ultrasound guidance to obtain a tissue biopsy. A smear of the aspirated material was prepared on a glass slide, air-dried and immediately stained with standard hematoxylin and eosin for rapid on-site evaluation (ROSE) by a cytopathologist. Three consecutive biopsies of a reactive LN were considered negative for malignant disease. The aspirated materials were also collected in liquid formalin, and the cell block underwent a final histological evaluation by the pathologist. The final pathological diagnosis was the standard reference for positive malignant LN infiltration. Patients with negative pathology
results on EBUS in whom malignant LNs were confirmed after surgery (mediastinoscopy or lung resection with systematic lymphadenectomy) were also considered positive in the sonographic analysis of the LNs.

**Analysis of ultrasound images**

All ultrasound images were evaluated to determine the foregoing sonographic features with the aim of validating a score predictive of malignancy. The following sonographic features associated with malignancy were evaluated:

1. Round shape, defined as a ratio of <1.5 between 2 perpendicular axes.
2. Distinct and well defined, distinguished by a marked white line delimiting the LN.
3. Appearance of heterogeneous echogenicity as a dichotomous variable (as opposed to homogeneous).
4. Absence of visible CHS in the form of a central linear structure with high echogenicity, with or without a blood vessel.
5. Small axis ≥10 mm.
6. Hyperechogenic density in the interior of the LN (Figs. 1 and 2).

The interrater agreement (IRA) was calculated for each sonographic criteria. These criteria were correlated with the final histopathological results, and the sensitivity, specificity and positive and negative predictive values (PPV and NPV, respectively) were calculated from the receiver operating characteristics (ROC) curve and the area under the curve (AUC).

**Scoring**

According to the reproducibility of the above-mentioned criteria, a simplified 5-item score was generated and calculated for prediction of LN malignancy. Each item was scored as follows: heterogeneous (1.5 points) versus homogeneous (0 points) echogenicity; absence of CHS (1.5 points) versus presence (0 points); round shape (1 versus 0 points); small LN axis ≥10 mm (1 point); and well-defined LN margin (1 point) versus indistinct margin (0 points). A total score was generated from the sum of each separate criterion and plotted relative to the pathological diagnosis on ROC curves.

**Score validation**

The previously modified score was applied prospectively to the ultrasound images of the suspect LN during the EBUS procedure by the two bronchoscopists performing the examination. The sonographic analysis was performed by a pathologist before the ROSE. The score was calculated and correlated with the histopathological results. A Doppler ultrasound was performed to evaluate blood flow within the LN. The utility of the proposed score was evaluated for the prediction of LN malignancy in real-time clinical application.

**Statistical analysis**

The data are presented as mean ± standard deviation (SD) or number (percentage), as appropriate. Malignant and non-malignant sonographic features were compared using the Mann–Whitney test. Statistical tests were performed with two-tailed p values, with p < 0.05 being considered statistically significant. ROC curves were used to calculate sensitivity and specificity, followed by the calculation of PPV and NPV for the score and for each criterion in relation to the final pathological diagnosis. AUC, C-index values and 95% confidence intervals (CI) are presented. The IRA was calculated using the interval-to-interval method, which is calculated by adding the total number of agreements in the intervals (for both the presence and the absence of malignancy) and dividing by the sum of the number of agreements and discrepancies.
in the intervals\textsuperscript{16}; an IRA of 80\% was considered a positive agreement. MedCalc\textsuperscript{\textregistered} (version 9.2.1.0, Acacialaan 22, B-8400 Ostend, Belgium) statistical software was used.

## Results

### Sonographic features

Patient characteristics and principal results are summarized in Table 1. After correlation with the pathology data was established, the images obtained from 141 patients, including 208 LNs from various node stations, were analyzed. The subcarinal LN (station 7) was the most frequently aspirated node (45.7\%), followed by stations 4R and 4L (24\% and 9.6\%, respectively) (Table 1).

A definitive diagnosis of malignant disease in the LNs analyzed was obtained in 50.5\% of the subjects compared to 49.5\% with reactive lymphadenitis. The final pathology of the biopsies evaluated is presented in Table 1, showing that adenocarcinoma was the most common diagnosis in the study cohort (22.7\%), followed by squamous cell carcinoma (12.8\%).

The analysis of the LN sonographic features, with their corresponding sensitivity, specificity, PPV and NPV for malignant disease, is shown in Table 2. The evaluation of the different sonographic criteria revealed that both heterogeneous echogenicity and the absence of CHS with or without central blood vessel were the most sensitive factors for predicting malignant disease (99\% in both cases), while hyperechogenic density in the interior of the LN had the highest specificity for predicting malignant LNs (78\%). Moreover, heterogeneous echogenicity showed an NPV of 94.4\%, followed by the absence of CHS (90\%). None of the proposed sonographic criteria had a clinically acceptable PPV. A positive IRA for the proposed criteria was recorded, with the maximum agreement being observed for heterogeneous echogenicity (92.2\%) (Table 2). The difference between the malignant and non-malignant infiltration of the LN for each criterion was statistically significant ($p < 0.01$), with the exception of the presence of hyperechogenic density in the interior of the LN ($p > 0.05$). Consequently, scoring was generated using only 5 criteria, as described above in the methodology section. With the ROC analysis, the combined score of $> 5$ has a sensitivity of 73.3\% and a specificity of 68\% (AUC = 0.738; 95\% CI: 0.673–0.796; $p = 0.0001$) for the prediction of malignant LN infiltration (Fig. 3). The PPV and NPV for the generated score were 70\% and 71.4\%, respectively.

### Table 1

<table>
<thead>
<tr>
<th>Age (years) [mean ± SD]</th>
<th>Males: $n$ (%)</th>
<th>Females: $n$ (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>63.1 ± 10</td>
<td>117 (83)/24 (17)</td>
</tr>
<tr>
<td>Affected LN ($n = 208$)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7, $n$ (%)</td>
<td>95 (45.7)</td>
<td></td>
</tr>
<tr>
<td>4R, $n$ (%)</td>
<td>50 (24)</td>
<td></td>
</tr>
<tr>
<td>4L, $n$ (%)</td>
<td>20 (9.6)</td>
<td></td>
</tr>
<tr>
<td>2R, $n$ (%)</td>
<td>8 (3.8)</td>
<td></td>
</tr>
<tr>
<td>6, $n$ (%)</td>
<td>1 (0.4)</td>
<td></td>
</tr>
<tr>
<td>Stations 10–12, $n$ (%)</td>
<td>34 (16.3)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Pathology diagnosis: malignant cases</th>
<th>Population ($n = 141$)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Primary lung cancer</strong></td>
<td></td>
</tr>
<tr>
<td>Adenocarcinoma, $n$ (%)</td>
<td>32 (22.7)</td>
</tr>
<tr>
<td>Squamous cell carcinoma, $n$ (%)</td>
<td>18 (12.8)</td>
</tr>
<tr>
<td>Small cell lung cancer, $n$ (%)</td>
<td>9 (6.4)</td>
</tr>
<tr>
<td>Undifferentiated non-small cell lung cancer, $n$ (%)</td>
<td>9 (6.4)</td>
</tr>
<tr>
<td>Small cell carcinoma, $n$ (%)</td>
<td>2 (1.4)</td>
</tr>
<tr>
<td><strong>Other</strong></td>
<td></td>
</tr>
<tr>
<td>Extrathoracic metastasis, $n$ (%)</td>
<td>8 (5.7)</td>
</tr>
<tr>
<td>Lymphoma, $n$ (%)</td>
<td>3 (2.1)</td>
</tr>
</tbody>
</table>

L, left; LN, lymph node; R, right; SD, standard deviation.
Table 2
Sensitivity, specificity, positive predictive value and negative predictive value for lymph node malignancy according to the different sonographic criteria.

<table>
<thead>
<tr>
<th>Sonographic criteria</th>
<th>Sensitivity (95% Cl)</th>
<th>Specificity (95% Cl)</th>
<th>PPV (95% Cl)</th>
<th>NPV (95% Cl)</th>
<th>IRA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Round shape</td>
<td>91.3% (83.8–95.7)</td>
<td>26.2% (18.3–36)</td>
<td>55.6% (47.8–63.1)</td>
<td>75% (57.5–87.3)</td>
<td>91%</td>
</tr>
<tr>
<td>Short axis ≥1 cm</td>
<td>89.5% (81.6–94.4)</td>
<td>36% (26.9–46)</td>
<td>69.1% (60.5–76.6)</td>
<td>72.5% (55.9–84.9)</td>
<td>86%</td>
</tr>
<tr>
<td>Heterogeneous echogenicity</td>
<td>99% (94–99.9)</td>
<td>16.5% (10.2–25.4)</td>
<td>54.7% (47.4–62)</td>
<td>94.4% (70.6–99.7)</td>
<td>92.2%</td>
</tr>
<tr>
<td>Hyperechogenic density in the interior of the LN</td>
<td>21% (13.9–30.2)</td>
<td>78% (68.4–85.2)</td>
<td>49% (34–64)</td>
<td>49.4% (41.5–57.3)</td>
<td>81.6%</td>
</tr>
<tr>
<td>Absence of any CHS</td>
<td>99% (94.8–99.8)</td>
<td>8.7% (4.1–16.3)</td>
<td>52.5% (43.5–59)</td>
<td>90% (54.1–99.5)</td>
<td>87.2%</td>
</tr>
<tr>
<td>Distinct margin</td>
<td>90.5% (82.8–95.1)</td>
<td>24.3% (16.6–33.9)</td>
<td>55% (47.2–62.4)</td>
<td>71.4% (53.5–84.8)</td>
<td>91.5%</td>
</tr>
</tbody>
</table>

CHS, central hilar structure; CI, confidence interval; IRA, interrater agreement; LN, lymph node; ND, not determined; NPV, negative predictive value; PPV, positive predictive value.

**Prospective validation study**

A real-time evaluation of the LNs during EBUS was performed in 39 patients, and 65 LNs were examined during the procedure. The prevalence of malignant disease in the LNs examined was 35.4% and the most common pathological diagnosis was adenocarcinoma (15.4%). The subcarinal LN was the most frequently aspirated station (46.2%), followed by station 4R (24.6%) and then station 10R or 10L (18.5%). Stations 4L, 2R and 2L represented 7.7%, 1.5% and 1.5% of the aspirations, respectively. When the previously validated sonographic criteria were applied in real time, the absence of CHS, a round shape and a short axis ≥10 mm yielded significantly statistical differences between non-malignant and malignant LNs (p = 0.0012; p = 0.012 and p = 0.023, respectively). Neither heterogeneous echogenicity nor a distinct margin showed a statistically significant difference between benign and malignant LNs (p > 0.05), as indicated in the retrospective analysis of the ultrasound images.

With the combination of positive criteria during the procedure, a score of >5 was found to have a sensitivity of 78% and specificity of 86% in the detection of a malignant LN (AUC = 0.852; 95% CI: 0.743–0.928; p = 0.0001), with a greater specificity than that obtained from the analysis of images. PPV and NPV were 75% and 88%, respectively.

**Discussion**

EBUS-TBNA has been welcomed in clinical practice as an effective instrument for lung cancer staging with an accuracy comparable to that of surgery. However, systematic aspiration of all LNs by EBUS-TBNA is a lengthy and expensive procedure. For this reason, the need for sonographic criteria to serve as a guide for obtaining biopsies from the various LN stations has been identified. This study has shown that no single criterion is sufficiently specific, but a combination of different sonographic criteria may be useful for predicting malignancy in mediastinal LNs and this combination of criteria may be valid for real-time clinical application.

Sonographic criteria for the prediction of malignant LN infiltration in lung cancer have already been analyzed retrospectively in various studies in EBUS or EUS. Fujiwara et al. defined 4 different criteria for predicting malignancy in LNs, with a diagnostic accuracy of between 63.8% and 86.0% for malignant LNs. In a similar study, Gill et al. explored mediastinal LNs in cases of primary lung cancer using EUS; these authors found that only a round shape, short axis >8.3 mm and distinct margins were associated with malignant LN infiltration. More recently, Schmid-Bindert et al. studied mediastinal lymphadenopathies, irrespective of underlying disease, and established a scoring system based on the sum of 6 positive criteria suggestive of LN malignancy; they described a high probability of malignant disease if the score was >3. They also found that heterogeneous echogenicity and the absence of CHS provided better NPV. These results coincide with the data presented here.

In this study, 5 criteria for LN malignancy were selected for their simplicity, making them suitable for use in real-time situations. Although none of them in isolation was sufficiently specific for differentiating malignancy, as was also found in earlier studies, a score of >5 was associated with acceptable accuracy in terms

**Fig. 3.** ROC curve of the score for the prediction of malignant lymph node infiltration. (A) ROC curve for the retrospective study showing different sensitivities and specificities of the combined positive criteria (AUC = 0.738; 95% CI: 0.673–0.796; p = 0.0001). (B) ROC curve for the prospective study (AUC = 0.852; 95% CI: 0.743–0.928; p = 0.0001).
of sensitivity and specificity in the retrospective analysis. In line with previously obtained results, heterogeneous echogenicity and the absence of CHS were the most sensitive factors in predicting LN malignancy (99%) and provided the best NPV (94.4% and 90%, respectively). It is interesting to observe that the real-time application of these criteria during the EBUS procedure improved both the sensitivity and the specificity of a score of >5 for predicting LN malignancy by 78.3% and 86%, respectively. Moreover, the absence of CHS, the shape and the diameter of the short axis ≥ 10 mm were significantly different for non-malignant and malignant infiltration. However, heterogeneous echogenicity and a distinct margin showed no significant correlation during real-time evaluation, compared to the previous image analysis observations.

A retrospective analysis of the ultrasound images showed that the presence of hyperechogenic density in the interior of the LN was the most specific sign of malignancy, but sensitivity was low, probably due to its low incidence. Consequently, this criterion was not included in the combined scoring system. The unacceptable PPV of current criteria reviewed may be due, in part, to the relatively low prevalence of malignant LN infiltration in our cohort. A score of >5, meanwhile, was found to have the greatest sensitivity and specificity, and this may have several explanations: firstly, there is no single acceptable sonographic criterion for predicting LN malignancy; secondly, no malignant lymphadenopathies, including reactive LN, with a diameter of ≥ 10 mm and the absence of CHS on ultrasound have been described, indicating that these two features are non-specific for malignant LN infiltration. This raised the problem, discussed by previous authors, of basing prediction of malignant LN infiltration on only three positive criteria. The coagulation necrosis sign was not observed in our retrospective analysis, and the grade of blood flow to the interior of the LN was not feasible for simplified application in a clinical setting, and only achieved moderate agreement among the different raters, and as such was not included in the criteria evaluated.

This study has certain limitations. Firstly, measurement of the LN axis is subjective and depends on manipulation of the EBUS transducer for obtaining the maximum LN diameters. Secondly, biopsies could not be obtained for all LNs detected during the EBUS procedure due to the proximity of blood vessels or difficulty in introducing the needle through the walls of the airway due to, for example, fibrosis or malignant infiltration of the wall itself. Thirdly, EUS was not used in this study. A combination of EUS-guided fine needle aspiration (EUS-FNA) and EBUS-TBNA is recommended for full evaluation of mediastinal LNs. This provides greater sensitivity and NPV than those obtained with each technique used alone. However, a recent randomized controlled trial looking at diagnostic accuracy has shown that the addition of EUS-FNA to EBUS-TBNA did not improve accuracy or sensitivity.

To conclude, the selected criteria (round shape, distinct margin, heterogeneous echogenicity, the absence of CHS, and short axis ≥ 10 mm) are reliable for real-time clinical application. A score of >5 showed good prediction for malignant LN and may be useful for taking decisions regarding LN sampling during the procedure. Moreover, there is no single criterion that can be used to rule out malignant infiltration of mediastinal or hilar LN; however, the absence of CHS, a round shape and an increased LN axis are suggestive of LN malignancy.

Conflict of interests

Hanaa Shafiek has received grants from the University of Alexandria and the Egyptian Ministry of Upper Education, as member of ParOwn (The Partnership and Ownership Initiative). The other authors have no conflict of interests.

Acknowledgements

The authors thank all the members of the Bronchoscopy Unit team for their contributions to the logistics of this study.

References


