Clinical utility of artificial intelligence–augmented endobronchial ultrasound elastography in lymph node staging for lung cancer

Yogita S. Patel, BSc,[a] Anthony A. Gatti, PhD,[b] Forough Farrokhyar, MPhil, PhD,[c] Feng Xie, PhD,[d] and Waël C. Hanna, MDCM, MBA[a,c]

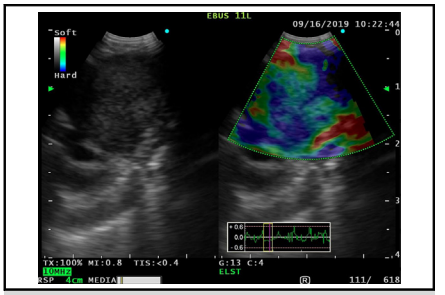
ABSTRACT

Objective: Endobronchial ultrasound elastography produces a color map of mediastinal lymph nodes, with the color blue (level 60) indicating stiffness. Our pilot study demonstrated that predominantly blue lymph nodes, with a stiffness area ratio greater than 0.496, are likely malignant. This large-scale study aims to validate this stiffness area ratio compared with pathology.

Methods: This is a single-center prospective clinical trial where B-mode ultrasound and endobronchial ultrasound elastography lymph node images were collected from patients undergoing endobronchial ultrasound transbronchial needle aspiration for suspected or diagnosed non–small cell lung cancer. Images were fed to a trained deep neural network algorithm (NeuralSeg), which segmented the lymph nodes, identified the percent of lymph node area above the color blue threshold of level 60, and assigned a malignant label to lymph nodes with a stiffness area ratio above 0.496. Diagnostic statistics and receiver operating characteristic analyses were conducted. NeuralSeg predictions were compared with pathology.

Results: B-mode ultrasound and endobronchial ultrasound elastography lymph node images (n = 210) were collected from 124 enrolled patients. Only lymph nodes with conclusive pathology results (n = 187) were analyzed. NeuralSeg was able to predict 98 of 143 true negatives and 34 of 44 true positives, resulting in an overall accuracy of 70.59% (95% CI, 63.50-77.01), sensitivity of 43.04% (95% CI, 31.94-54.67), specificity of 90.74% (95% CI, 83.63-95.47), positive predictive value of 77.27% (95% CI, 64.13-86.60), negative predictive value of 68.53% (95% CI, 64.05-72.70), and area under the curve of 0.820 (95% CI, 0.758-0.883).

Conclusions: NeuralSeg was able to predict nodal malignancy based on endobronchial ultrasound elastography lymph node images with high area under the receiver operating characteristic curve and specificity. This technology should be refined further by testing its validity and applicability through a larger dataset in a multicenter trial. (JTCVS Techniques 2024;27:158-66)



Static B-mode ultrasound image (*left*) and EBUS elastography image (*right*) of an LN.

CENTRAL MESSAGE

NeuralSeg, an AI algorithm, was able to predict nodal malignancy directly from EBUS elastography LN images with high area under the ROC curve and high specificity.

PERSPECTIVE

NeuralSeg, an AI algorithm, was able to predict nodal malignancy directly from EBUS elastography LN images based on a predefined SAR cutoff of 0.496 above a color blue stiffness threshold of level 60 with high area under the ROC curve and specificity. This study is a step forward in the applicability of AI in detecting mediastinal LN malignancy.

From the [a]Division of Thoracic Surgery, Department of Surgery, [c]Department of Health Research Methods, Evidence and Impact, [d]Department of Epidemiology and Biostatistics, McMaster University, Hamilton, Ontario, Canada; and [b]Department of Radiology, Stanford University, Stanford, Calif.

No funding was obtained for this study. Olympus Corporation of the Americas, Center Valley, Pennsylvania, provided supplies of the Olympus EU-ME2 Plus Ultrasound Transducer with Elastography Module and related accessories as in-kind donation for the duration of this study.

Institutional Review Board Approval: Hamilton Integrated Research Ethics Board, August 4, 2021, Project #12644. Informed patient consent was obtained.

Read at the Society of Thoracic Surgeons 59th Annual Meeting, San Diego, California, January 21-23, 2023.

Received for publication March 20, 2024; revisions received May 28, 2024; accepted for publication June 17, 2024; available ahead of print July 24, 2024.

Address for reprints: Waël C. Hanna, MDCM, MBA, Division of Thoracic Surgery, Department of Surgery, St Joseph's Healthcare Hamilton, 50 Charlton Ave E, Suite T-2105F, Hamilton, Ontario, L8N 4A6 Canada (E-mail: hannaw@mcmaster.ca).

2666-2507

https://doi.org/10.1016/j.xjtc.2024.06.024

Abbreviations and Acronyms

AI	= artificial intelligence
AUC	= area under the curve
CLNS	= Canada Lymph Node Score
EBUS-TBNA	= endobronchial ultrasound transbronchial needle aspiration
IQR	= interquartile range
LN	= lymph node
NPV	= negative predictive value
PPV	= positive predictive value
ROC	= receiver operating characteristic
ROI	= region of interest
SAR	= stiffness area ratio

The staging of mediastinal lymph nodes (LNs) is a crucial step in the lung cancer diagnostic pathway.[1,2] The current guideline for mediastinal nodal staging for non–small cell lung cancer is endobronchial ultrasound transbronchial needle aspiration (EBUS-TBNA).[3,4] Despite advances in staging procedures, recent studies show that approximately 40% to 74% of mediastinal metastases are diagnosed correctly via EBUS,[5-7] diagnostic yield is 64%,[8] sample adequacy with EBUS is approximately 70%,[9] and as high as 40% of biopsies result in inconclusive pathology,[10] and this is because EBUS-TBNA is highly operator dependent and factors such as the skill of the endoscopist and cytologists impact the accuracy of EBUS-TBNA.[11] This high rate of inconclusive results may lead to repeated or additional investigations and perhaps delay definitive treatment. It is possible that the complexity of EBUS-TBNA, coupled with a high rate of inconclusive results, is a contributing factor to why some studies show that as little as 22% to 27% of patients at risk for nodal disease on imaging receive preoperative nodal staging in the United States.[12-15] Therefore, an adjunct to EBUS-TBNA for mediastinal LN staging is warranted.

Elastography is a novel technology that can produce visual color maps representative of tissue stiffness, where red represents soft tissue, and blue represents stiff tissue.[16] Elastography can be applied to mediastinal LN staging because malignant tissue tends to be stiffer in nature because there are more cells per area compared with benign.[17] A systematic review and meta-analysis by Wu and colleagues[18] in 2022 evaluated EBUS elastography in differentiating benign and malignant mediastinal and hilar LNs. This study included 2307 LNs from 17 studies and obtained a pooled sensitivity of 0.90 (95% CI, 0.84-0.94), pooled specificity of 0.78 (95% CI, 0.74-0.81), and area under the curve (AUC) of 0.86 (95% CI, 0.82-0.88).[18] Different qualitative and quantitative methods were used in the included studies to analyze the elastography color maps.[18] Qualitative methods include 3- or 5-type classification methods to observe and categorize images[19,20]; however, these methods are often subjective and dependent on the interpreter. Quantitative methods include strain ratio, stiffness area ratio (SAR), and strain histogram, which tend to be far more reproducible than qualitative methods.[16,21,22] Of these methods, the most intuitive is the SAR, because it is calculated based on the number of blue pixels compared with all the color pixels in the region of interest (ROI), the LN.[16] However, this method has yet to become standardized because both the stiffness threshold of the color blue and the SAR cutoffs have not been defined.

In our previous pilot study, we used a trained deep neural network artificial intelligence (AI) algorithm (NeuralSeg) to segment 31 EBUS elastography LN images to define the color blue stiffness threshold required to calculate the SAR and determine the optimal SAR to distinguish between benign and malignant LNs.[23] Nine predefined color blue stiffness thresholds were tested, and a color blue stiffness threshold of level 60 from the 0-255 color scale was found to have the highest AUC of 0.891.[23] An optimal SAR cutoff of 0.496 was also determined for predicting LN malignancy, with an overall accuracy of 83.30%, sensitivity of 92.30%, and specificity of 76.50%.[23] However, this SAR remains difficult to determine with the human eye. On a complex EBUS elastography color map of an LN, it is practically impossible for the endoscopist to determine the SAR by visual inspection alone. We believe that the incorporation of AI into the ultrasound system may allow real-time determination of the SAR and prediction of LN malignancy. As a first step, we hypothesized that NeuralSeg could predict LN malignancy from EBUS elastography images of mediastinal LNs based on the SAR cutoff of 0.496 above a color blue stiffness threshold of level 60. The results were compared with the gold standard of final pathology results from surgical specimens (in patients who underwent resection) or diagnostic nodal biopsies (in patients who did not undergo surgery) to determine the diagnostic capability of AI-augmented EBUS elastography.

MATERIALS AND METHODS

Study Design

This is a single-center, prospective clinical trial where B-mode ultrasound and EBUS elastography LN images were collected at the time of the EBUS-TBNA procedures.

Research Ethics Approval and Trial Registration

This study was reviewed and approved on August 4, 2021, by the Hamilton Integrated Research Ethics Board, Project #12644. This trial was registered on ClinicalTrials.gov (#NCT04816981).

Study Subjects

Patients undergoing EBUS-TBNA for mediastinal LN staging for confirmed or suspected non–small cell lung cancer between August 2021

and May 2022 at St Joseph's Healthcare Hamilton, Hamilton, Ontario, Canada, were eligible for this study. No exclusion criteria were applied. Patients provided consent before the EBUS-TBNA procedure, and enrollment in the study did not intervene with the standard of care practices. Patients were enrolled in a consecutive sample, and patient involvement concluded when the procedure ended. No follow-up was required after the procedure.

Endobronchial Ultrasound Transbronchial Needle Aspiration Procedure

An Olympus convex probe ultrasound bronchoscope (BF-UC180F; Olympus) and the EU-ME2 plus transducer (Olympus) were used to perform EBUS-TBNA under conscious sedation with midazolam and fentanyl. An expert endosonographer assigned a Canada Lymph Node Score (CLNS), a 4-point score that can be used during EBUS-TBNA to help identify LN malignancy, to each mediastinal LN based on the 4 ultrasonographic nodal features that are predictive of LN malignancy: short-axis diameter 10 mm or greater, presence of well-defined margins, absence of the central hilar structure, and presence of central necrosis[24] before conducting the biopsy by TBNA using a 22-gauge needle (NA-201SX-4022, Olympus). Specimen adequacy was confirmed by rapid onsite cytology.

Endobronchial Ultrasound Elastography

Before LN biopsy, EBUS elastography was performed. The ROI was the LN, and a 1:1 ratio with the surrounding mediastinal tissue was ensured. The strain graph was used to confirm stable pressurization by ensuring the wave was between −0.6 and +0.6 (Figure 1, *A*). The B-mode and EBUS elastography images of the LN were displayed side-by-side, and a suitable static image (Figure 1, *B* and *C*) was captured and stored on an external hard drive as Red Green Blue Joint Photographic Experts Group images.

Data Collection

Patient demographics, including age, gender, and smoking status, were collected. LN characteristics, including LN station, short- and long-axis measurements from computed tomography scans, standardized uptake values from positron emission tomography scans, the CLNS, and final pathology results via the immunohistochemistry technique, were also obtained.

Unit of Analysis

The unit of analysis for this study was the LN rather than the patient, because the primary outcome is whether NeuralSeg could predict LN malignancy directly from the EBUS elastography image of a mediastinal LN based on the SAR cutoff of 0.496 above a color blue stiffness threshold of level 60.[23]

Sample Size

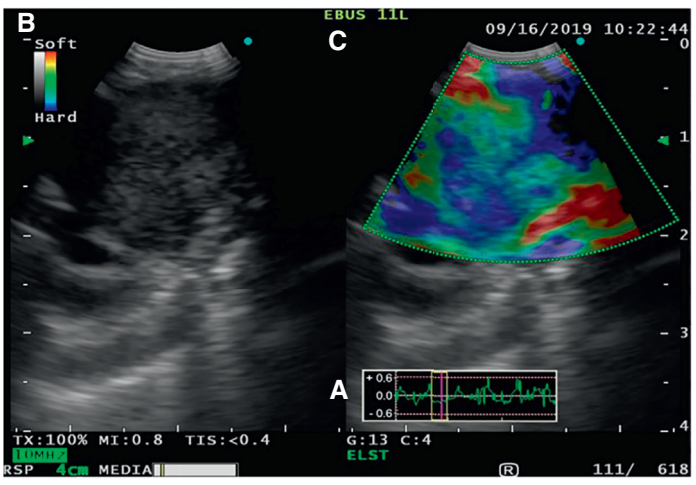
Elastography for LN imaging is associated with a sensitivity of 92.30% for SAR based on our pilot study.[23] Assuming this diagnostic value, and a prevalence of 18% malignancy in the lung cancer population at our center, to reach 80% power, an alpha of 0.05, and a marginal error of 0.085 with a 95% CI for sensitivity, a sample size of 210 LNs was needed.[25]

Outcomes

The primary outcome was to test whether NeuralSeg, a deep learning-based neural network AI algorithm, can predict LN malignancy directly from the EBUS elastography image of a mediastinal LN based on the SAR cutoff of 0.496 above a color blue stiffness threshold of level 60 to validate the findings from our pilot study.[23] The gold standard for comparison was the final pathology results from surgical specimen (in patients who underwent resection) or diagnostic nodal biopsies (in patients who did not undergo surgery) to determine the diagnostic capability of AI-augmented EBUS elastography.

Endobronchial Ultrasound Elastography Lymph Node Image Analysis

NeuralSeg, an AI deep neural network, was used to analyze the EBUS elastography LN images. NeuralSeg was previously trained and validated to segment LNs and extract the 4 ultrasonographic LN features predictive of malignancy (short-axis diameter ≥10 mm, well-defined margins, absence of hilar structure, and presence of central necrosis), based on the CLNS.[26] NeuralSeg was also previously trained in our pilot study to segment the LN from the B-mode image using the LN features listed above, overlay the segmentation onto the EBUS elastography image, identify the percent of LN area above the color blue stiffness threshold of level 60, and automatically assign a malignant or benign label based on the SAR cutoff of 0.496.[23]



D

Accuracy (95% CI)	70.59% (63.50% to 77.01%)
Sensitivity (95% CI)	43.04% (31.94% to 54.67%)
Specificity (95% CI)	90.74% (83.63% to 95.47%)
PPV (95% CI)	77.27% (64.13% to 86.60%)
NPV (95% CI)	68.53% (64.05% to 72.70%)

FIGURE 1. A, The strain graph from an EBUS elastography image, which is used to confirm standard pressurization by ensuring the wave is between −0.6 and +0.6. B, Static B-mode ultrasound image of LN and C, static EBUS elastography image of a LN captured during an EBUS-TBNA procedure, displayed side by side. D, Diagnostic statistics based on the SAR cutoff of 49.59% above the color blue stiffness threshold of level 60. *PPV*, Positive predictive value; *NPV*, negative predictive value.

In this study, static images of the B-mode and EBUS elastography images of the LNs were fed to NeuralSeg. The EBUS elastography images were separated from the B-mode images to isolate the color map and to minimize the effects of the B-mode image on the SAR analysis. The computer programmer training NeuralSeg was blinded to all data, including patient demographics, LN characteristics, and LN pathology. NeuralSeg automatically segmented the LN from the surrounding tissue on the B-mode side of the image using the LN features listed above and then overlaid the segmentation onto the EBUS elastography side of the image. NeuralSeg automatically identified the percent of LN area above the previously defined color blue stiffness threshold of level 60 and assigned a malignant or benign label based on the previously defined SAR cutoff of 0.496.[23] By using a 5-fold cross-validation technique derived from a previous study, 5 SARs were obtained for the color blue stiffness threshold of level 60. The mean of these 5 predictions was used as the final SAR. This was done to improve the robustness of the results.

Statistical Analyses

Descriptive statistics are provided for patient demographics and LN characteristics, categorical variables were reported as counts (percentages), and continuous variables as means [SDs] if they were normally distributed or as medians (interquartile range [IQR]) if they were not normally distributed. Descriptive statistics for SAR were compared between the benign and malignant LNs, continuous parametric variables were compared using the Student *t* test, and nonparametric variables were compared using Mann–Whitney *U* test. Diagnostic ability of AI-augmented EBUS elastography was assessed by obtaining diagnostic accuracy, sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and AUC. Receiver operating characteristic (ROC) curve analysis was used to determine optimal cutoff values to distinguish malignant and benign LNs. Backward stepwise logistic regression was performed to determine predictors of malignant LNs using short- and long-axis measurements, CLNS, and SAR. All statistical tests used 2-sided hypotheses. All statistical analyses were done using the 2020 version of Statistical Package for the Social Sciences (SPSS Inc).[27]

RESULTS

Patients and Lymph Nodes

Between August 2021 and May 2022, a total of 124 patients were enrolled, and a total of 210 B-mode ultrasound and EBUS elastography images of mediastinal LNs were collected (Figure 2). Patient demographics and LN characteristics are shown in Table 1. The mean (SD) age of the patients was 69.83 ± 9.95 years, 51.61% (64/124) were male, and mean (SD) body mass index was 26.69 ± 6.14 kg/m^2. In terms of smoking status, 35.48% (44/124) were current smokers, 43.55% (54/124) were former smokers, 8.06% (10/124) of patients had never smoked, and for 12.90% (16/124), the smoking status was unknown. Common comorbidities among the patients included hypertension (41.13%; 51/124), chronic obstructive pulmonary disease (24.19%; 30/124), diabetes (20.97%; 26/124), atrial fibrillation (9.68%; 12/124), and gastroesophageal reflux disease (8.06%; 10/124). Approximately half (51.43% [108/210]) of the LNs analyzed were benign, 37.62% (79/210) were malignant, 8.10% (17/210) had inconclusive results on pathology, and the pathology could not be obtained for 2.86% (6/210). The majority, 37.62% (79/210) of the LNs were obtained from station 7, 32.86% (69/210) were from station 4R, and 15.71% (33/210) were from station 4L. The mean (SD) LN short- and long-axis measurements were 11.00 ± 5.76 mm and 15.70 ± 6.00 mm, respectively. The overall CLNS was 0 for 36.67% (77/210) of LNs, 1 for 18.10% (38/210) of LNs, 2 for 11.90% (25/210) of LNs, 3 for 21.90% (46/210) of LNs, 4 for 10.48% (22/210) of LNs, and 0.95% (2/210) of LNs were missing a CLNS.

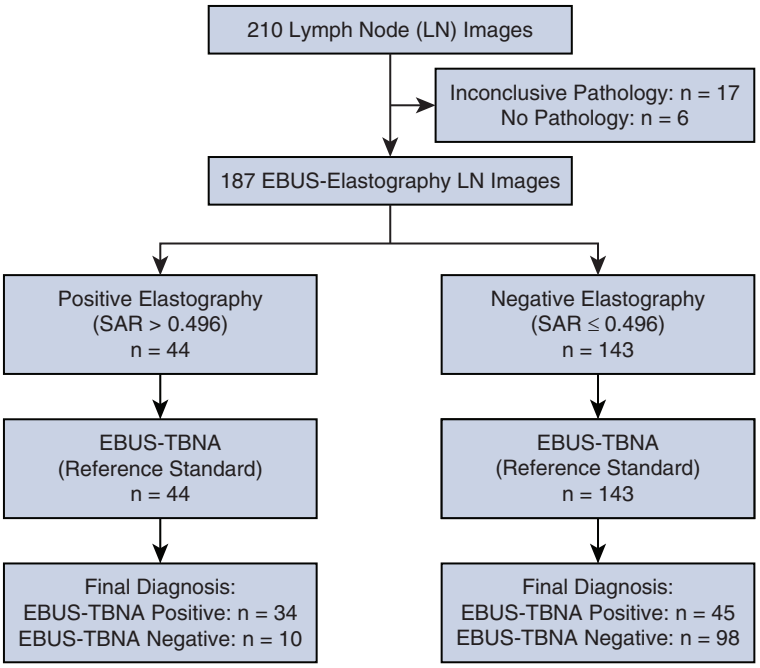


FIGURE 2. Flow chart of NeuralSeg's predictions directly from EBUS elastography images of LNs based on the SAR cutoff of 0.496 above a color blue stiffness threshold of level 60, compared with final pathology. *EBUS*, Endobronchial ultrasound; *SAR*, stiffness area ratio; *TBNA*, transbronchial needle aspiration.

TABLE 1. Patient demographics and lymph node characteristics

Variable	Patients (n = 124)
Age, y, mean ± SD	69.83 ± 9.95
Males, n (%)	64 (51.61)
Smoking status, n (%)	
Smoker	44 (35.48)
Former smoker	54 (43.55)
Never smoked	10 (8.06)
Unknown	16 (12.90)
BMI, kg/m^2, mean ± SD	26.69 ± 6.14
Comorbidities, n (%)	
Hypertension	51 (41.13)
Chronic obstructive pulmonary disease	30 (24.19)
Diabetes	26 (20.97)
Atrial fibrillation	12 (9.68)
Gastroesophageal reflux disease	10 (8.06)
Past cancer, n (%)	
Yes	29 (23.39)
No	90 (72.58)
Unknown	5 (4.03)
Variable	**LNs (n = 210)**
LN pathology, n (%)	
Benign	108 (51.43)
Malignant	79 (37.62)
Inconclusive	17 (8.10)
No pathology	6 (2.86)
LN station, n (%)	
7	79 (37.62)
4R	69 (32.86)
4L	33 (15.71)
8	7 (3.33)
10	2 (0.95)
11	3 (1.43)
Other	17 (8.10)
Short-axis measurement, mm, mean ± SD	11.00 ± 5.76
Long-axis measurement, mm, mean ± SD	15.70 ± 6.00
Ultrasound malignancy features based on CLNS, n (%)	
Short-axis diameter (≥10 mm)	98 (46.67)
Well-defined margins	102 (48.57)
Central hilar structure (absent)	71 (33.81)
Central necrosis (present)	44 (20.95)
CLNS, n (%)	
0	77 (36.67)
1	38 (18.10)
2	25 (11.90)
3	46 (21.90)
4	22 (10.48)
Unknown	2 (0.95)

BMI, Body mass index; *LN*, lymph node; *CLNS*, Canada Lymph Node Score.

The characteristics of the benign and malignant LNs are shown in Table 2. There were no significant differences in LN station between the benign and malignant LNs; however, the short- and long-axis measurements, ultrasound malignancy features based on the CLNS (short-axis diameter 10 mm or greater, well-defined margins, the absence of central hilar structure, and the presence of central necrosis), and the CLNS were significantly different between the benign and malignant LNs.

Validation of the Stiffness Area Ratio Cutoff

NeuralSeg was fed only the benign and malignant images (n = 187), and images with inconclusive or no pathology were removed from the dataset, because otherwise, no ground truth data would be available (Figure 2). Based on the SAR cutoff of 0.496 above the color blue stiffness threshold of level 60, NeuralSeg automatically assigned those LNs with a SAR greater than 0.496 a malignant label and those with a SAR equal to or below 0.496 a benign label. These results were compared with the biopsy or surgical pathology results. NeuralSeg was able to predict 68.53% (98/143) of true negatives and 77.27% (34/44) of true positives. This resulted in an overall diagnostic accuracy of 70.59% (95% CI, 63.50-77.01), sensitivity of 43.04% (95% CI, 31.94-54.67), specificity of 90.74% (95% CI, 83.63-95.47), PPV of 77.27% (95% CI, 64.13-86.60), and NPV of 68.53% (95% CI, 64.05-72.70) (Figure 1, *D*). The ROC curve for the color blue stiffness threshold of level 60 is shown in Figure 3, showing an AUC of 0.820 (95% CI, 0.758-0.883). When compared with final pathology results, the mean (SD) and median (IQR) SARs above the color blue stiffness threshold of level 60 were significantly different between the benign (0.25 ± 0.17; 0.21 [IQR, 0.14-0.32]) and malignant LNs (0.45 ± 0.15; 0.46 [IQR, 0.33-0.54]; $P < .001$). The results are shown in Table 3 and Figure 4. Backward stepwise logistic regression revealed that the SAR (odds ratio [OR], 51.69, 95% CI, 4.06-657.52; $P = .002$) and CLNS (OR, 2.92, 95% CI, 2.06-4.12; $P < .001$) were predictors of malignant LNs.

DISCUSSION

This study demonstrates that NeuralSeg, a deep neural network AI algorithm, could predict LN malignancy from EBUS elastography images of mediastinal LNs by successfully calculating SAR of the LNs above a color blue stiffness threshold of level 60 and automatically assigning a benign or malignant label based on the SAR cutoff of 0.496. It does so with high area under the ROC curve (consistent with an excellent diagnostic test)[28,29] and high

TABLE 2. Characteristics of benign and malignant lymph nodes

Characteristic	Benign LNs (n = 108)	Malignant LNs (n = 79)	*P* value
LN station, n (%)			
7	41 (37.96)	32 (40.51)	.21
4R	38 (35.19)	21 (26.58)	
4L	20 (18.52)	9 (11.39)	
8	3 (2.78)	4 (5.06)	
10	-	2 (2.53)	
11	-	3 (3.80)	
Other	6 (5.56)	8 (10.13)	
Short-axis measurement, mm, mean ± SD	8.88 ± 4.21	14.71 ± 6.13	**<.001**
Long-axis measurement, mm, mean ± SD	14.09 ± 4.51	19.03 ± 6.74	**<.001**
Ultrasound malignancy features based on CLNS, n (%)			
Short-axis diameter (≥10 mm)	30 (28.30)	63 (80.77)	**<.001**
Well-defined margins	33 (30.84)	63 (80.77)	**<.001**
Central hilar structure (absent)	13 (12.15)	52 (66.67)	**<.001**
Central necrosis (present)	7 (6.54)	34 (43.59)	**<.001**
CLNS, n (%)			
0	58 (53.70)	5 (6.33)	**<.001**
1	27 (25.00)	7 (8.86)	
2	12 (11.11)	11 (13.92)	
3	8 (7.41)	38 (48.10)	
4	2 (1.85)	17 (21.52)	
Unknown	1 (0.93)	1 (1.27)	

Bold shows statistically significant *P* values. *CLNS*, Canada Lymph Node Score; *LN*, lymph node.

specificity, but with low sensitivity. When compared with final pathology results, the mean SARs above the color blue stiffness threshold of level 60 were significantly different between benign (0.25) and malignant LNs (0.45; $P < .001$). This demonstrates that the color blue stiffness threshold of level 60 can accurately distinguish between benign and malignant LNs, and is a promising step in standardizing the color blue stiffness threshold to compute the SAR, thus improving the clinical utility of EBUS elastography in diagnosing mediastinal LNs. This adjunct to ultrasound examination also may help guide the decisions around whether repeat procedures are necessary in the instances when biopsy at the time of EBUS-TBNA is not feasible, such as when a patient is not tolerating the procedure, the LN is too small, or multiple passes result in a nondiagnostic specimen.

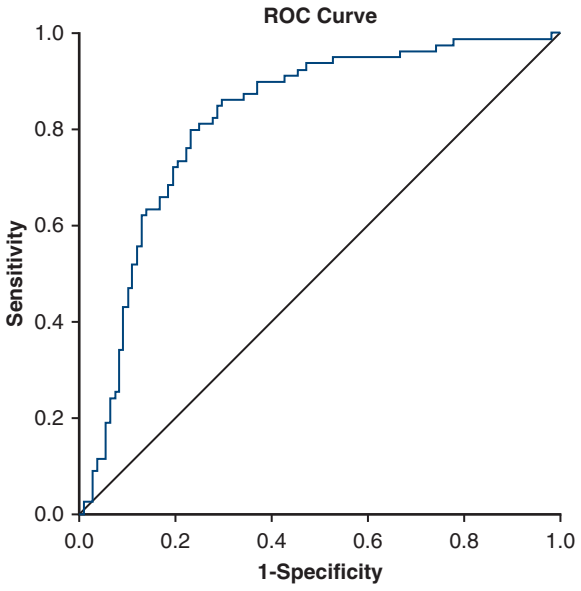


FIGURE 3. The ROC curve for the optimal color blue stiffness threshold of level 60. *ROC*, Receiver operating characteristic.

Five other studies looked at SAR in mediastinal LNs; however, each used human input with the ImageJ image analysis software (National Institutes of Health) or Photoshop (Adobe Systems Inc) to manually select the ROI and calculate SAR.[16,30-33] The results of these studies are as follows: Wang and colleagues[30] conducted a single-center retrospective study in 2023 with 131 LNs from 83 patients and determined a SAR cutoff of 0.60 with diagnostic accuracy of 84.4%, sensitivity of 83.3%, specificity of 86.0%, PPV of 89.6%, NPV of 78.2%, and AUC of 0.875. Uchimura and colleagues[31] conducted a single-center retrospective study in 2020 with 149 LNs from 132 patients and determined a SAR cutoff of 0.41 with a diagnostic accuracy of 83.9%, sensitivity of 88.2%, specificity of 80.2%, PPV of 78.9%, NPV of 89.0%, and AUC of 0.884. Fujiwara and colleagues[32] conducted a single-center retrospective study in 2019 with 228 LNs from 122 patients and determined a SAR cutoff of 0.31 with a diagnostic accuracy of 79.7%,

TABLE 3. Descriptive statistics for stiffness area ratio for benign and malignant lymph nodes

Pathology	n	Minimum	Maximum	Mean*	SD	Median†	IQR
Benign	108	0.016	0.81	0.25	0.17	0.21	0.14-0.32
Malignant	79	0.037	0.78	0.45	0.15	0.46	0.33-0.54

IQR, Interquartile range. *$P < .001$. *P* value obtained using the Mann–Whitney *U* test for continuous variable. †$P < .001$. *P* value obtained using the Wilcoxon rank-sum test.

sensitivity of 72.1%, specificity of 84.0%, PPV of 72.1%, and NPV of 84.0%. Ma and colleagues[33] conducted a single-center prospective study in 2018 with 79 LNs from 60 patients and determined a SAR cutoff of 0.367 with a diagnostic accuracy of 78.5%, sensitivity of 92.3%, specificity of 67.5%, and AUC of 0.86. Nakajima and colleagues[16] conducted a single-center retrospective study in 2015 with 49 LNs from 21 patients and determined a SAR cutoff of 0.311 with a sensitivity of 81%, and specificity of 85%. Although the sensitivity of our AI algorithm is lower than what is found in the existing literature, the specificity of our AI algorithm is much higher, indicating that our AI algorithm can effectively rule out nodal metastases.

AI, specifically deep neural networks, is a technology that is gaining popularity in medicine, especially for predicting and diagnosing diseases based on medical images.[34,35] Recent research has shown that AI and deep learning are being used to accurately interpret images when compared with clinicians in radiology, pathology, and cardiology.[36] The more data these AI algorithms are exposed to, the more they learn. Furthermore, using AI reduces the rater variability and increases the potential of this becoming a reproducible method in detecting LN malignancy. However, based on our study results, NeuralSeg could be optimized with further training and more data to improve its diagnostic capabilities. Nonetheless, this study is valuable because it demonstrates the feasibility of NeuralSeg, a deep neural network AI algorithm, and its ability to automatically segment mediastinal LNs based on ultrasonographic features and predict malignancy based on the predetermined SAR cutoff.

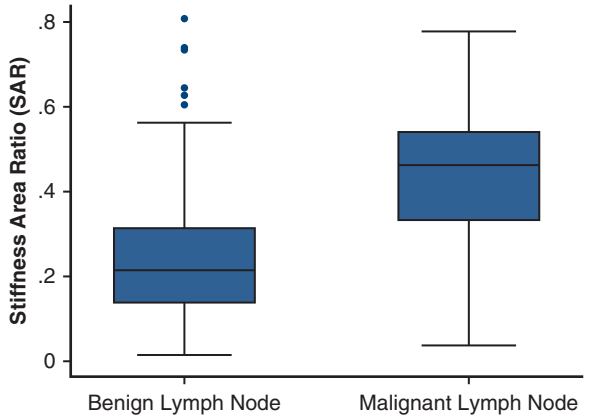


FIGURE 4. Box plot showing the SAR of the benign and malignant LNs. The lower and upper borders of the box represent the lower and upper quartiles (25th percentile and 75th percentile). The *middle horizontal line* represents the median. The lower and upper whiskers represent the minimum and maximum values of nonoutliers. The extra *dots* represent outliers. When compared with final pathology results, the median [IQR] SARs above the color blue stiffness threshold of level 60 were significantly different between the benign (0.21 [IQR, 0.14-0.32]) and malignant LNs (0.46 [IQR, 0.33-0.54]; $P < .001$).

Study Limitations

This study is not without limitations. First, this is the experience from a single-site, single-endosonographer study. Therefore, a multicenter, prospective trial with a larger sample size is required to perform external validation to further verify the results and confirm the reliability of the AI algorithm. Also, there was a relatively low prevalence of malignant LNs in this study. Additionally, only 4 ultrasonographic features (LN small-axis diameter, central hilar structure, central necrosis, and margin status) were used to train the algorithm. Although these features are predictive of malignancy, other ultrasonographic features may contribute to NeuralSeg's ability to predict malignancy. Another limitation is that only 1 elastography parameter was used, the SAR. Although SAR is an intuitive quantitative method that has achieved good diagnostic results, elastography is still a relatively new technology; therefore, all quantitative and qualitative methods should be thoroughly analyzed. Last, the quality of the static EBUS elastography images is not consistent. Different factors, such as stable pressurization and the amount of the ROI in the elastography frame, could potentially impact the results. Therefore, future studies should consider the strain wave and ROI to improve the EBUS elastography image quality to obtain the best possible EBUS elastography images.

CONCLUSIONS

In this study, NeuralSeg, a deep neural network AI algorithm, was able to predict nodal malignancy directly from EBUS elastography LN images based on a predefined SAR cutoff of 0.496 above a color blue stiffness threshold of level 60 with high area under the ROC curve and high specificity (Figure 5). This study is a meaningful step forward in the applicability of AI in detecting mediastinal LN malignancy at bedside and in real-time. However, more extensive multicenter studies must be conducted to standardize this process and optimize the AI algorithm (Figure 5).

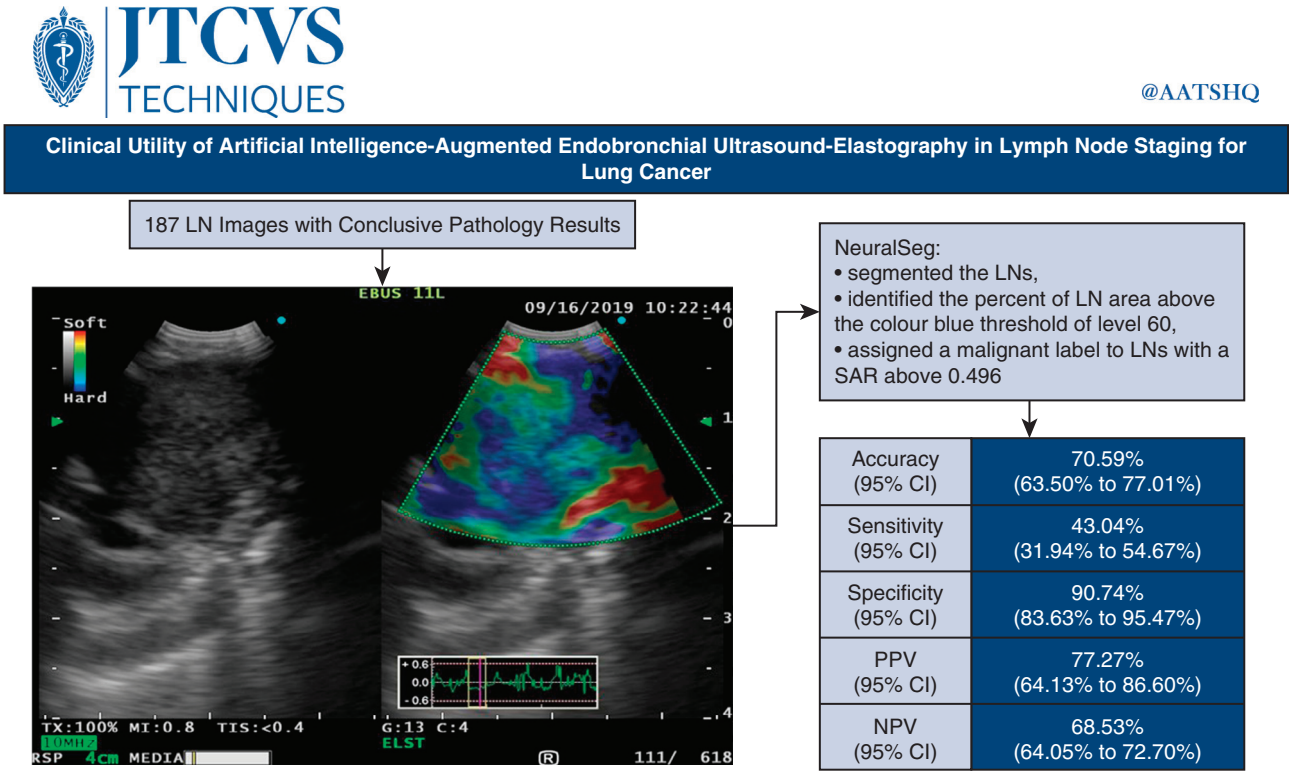


FIGURE 5. Graphical Abstract. *LN*, Lymph node; *SAR*, stiffness area ratio; *PPV*, positive predictive value; *NPV*, negative predictive value; *EBUS*, endobronchial ultrasound.

Conflict of Interest Statement

A.A.G.: ACLIP LLC (Owner/Co-Owner Founder/Co-Founder, Ownership Interest (stocks, stock options, patent or other intellectual property or other ownership interest excluding diversified mutual funds)), NeuralSeg Ltd (Owner/Co-Owner Founder/Co-Founder). W.C.H.: Hamilton Academic Health Sciences Organization AFP Innovation Grant, Advisory Board and Speakers Bureau AstraZeneca, Data Safety Monitoring Committee for Roche/Genentech, Speakers Bureau for Minogue Medical, and Grant Funding from Intuitive Surgical. All other authors reported no conflicts of interest.

The *Journal* policy requires editors and reviewers to disclose conflicts of interest and to decline handling or reviewing manuscripts for which they may have a conflict of interest. The editors and reviewers of this article have no conflicts of interest.

The authors thank Olympus Corporation of the Americas, Center Valley, Pennsylvania, for providing supplies of the Olympus EU-ME2 Plus Ultrasound Transducer with Elastography Module and related accessories as in-kind donation for the duration of this study. Honorable mention to Nikkita Mistry, MSc, for her assistance.

References

1. Leiro-Fernández V, Fernández-Villar A. Mediastinal staging for non-small cell lung cancer. *Transl Lung Cancer Res*. 2021;10(1):496-505. https://doi.org/10.21037/tlcr.2020.03.08
2. De Leyn P, Dooms C, Kuzdzal J, et al. Revised ESTS guidelines for preoperative mediastinal lymph node staging for non-small-cell lung cancer. *Eur J Cardiothorac Surg*. 2014;45(5):787-798. https://doi.org/10.1093/ejcts/ezu028
3. Silvestri GA, Gonzalez AV, Jantz MA, et al. Methods for staging non-small cell lung cancer: diagnosis and management of lung cancer, 3rd ed: American College of Chest Physicians evidence-based clinical practice guidelines. *Chest*. 2013;143(5 Suppl):e211S-e250S. https://doi.org/10.1378/chest.12-2355
4. Sampsonas F, Kakoullis L, Lykouras D, Karkoulias K, Spiropoulos K. EBUS: faster, cheaper and most effective in lung cancer staging. *Int J Clin Pract*. 2018;72(2). https://doi.org/10.1111/ijcp.13053
5. Kim BG, Cho JH, Shin SH, et al. Diagnostic performance of endosonography to detect mediastinal lymph node metastasis in patients with radiological N1 non-small cell lung cancer. *Cancer Res Treat*. 2023;55(3):832-840. https://doi.org/10.4143/crt.2022.1428
6. Vial MR, O'Connell OJ, Grosu HB, et al. Diagnostic performance of endobronchial ultrasound-guided mediastinal lymph node sampling in early stage non-small cell lung cancer: a prospective study. *Respirology*. 2018;23(1):76-81. https://doi.org/10.1111/resp.13162
7. Badaoui A, De Wergifosse M, Rondelet B, et al. Improved accuracy and sensitivity in diagnosis and staging of lung cancer with systematic and combined

endobronchial and endoscopic ultrasound (EBUS-EUS): experience from a tertiary center. *Cancers (Basel)*. 2024;16(4):728. https://doi.org/10.3390/cancers16040728
8. Demirkol B, Tanrıverdi E, Gül Ş, et al. The role of endobronchial ultrasonography elastography in the diagnosis of hilar and mediastinal lymph nodes. *Turk J Med Sci*. 2023;53(3):712-720. https://doi.org/10.55730/1300-0144.5634
9. Conte SC, Spagnol G, Biolo M, Confalonieri M. A retrospective study of endobronchial ultrasound transbronchial needle aspiration versus conventional transbronchial needle aspiration in diagnosis/staging of hilar/mediastinal lymph node in lung cancer: which role in clinical practice? *Monaldi Arch Chest Dis*. 2019; 89(1) https://doi.org/10.4081/monaldi.2019.1010
10. Ortakoylu MG, Iliaz S, Bahadir A, et al. Diagnostic value of endobronchial ultrasound-guided transbronchial needle aspiration in various lung diseases. *J Bras Pneumol*. 2015;41(5):410-414. https://doi.org/10.1590/S1806-37132015000004493
11. Jalil BA, Yasufuku K, Khan AM. Uses, limitations, and complications of endobronchial ultrasound. *Proc (Bayl Univ Med Cent)*. 2015;28(3):325-330. https://doi.org/10.1080/08998280.2015.11929263
12. Osarogiagbon RU, Lee YS, Faris NR, Ray MA, Ojeabulu PO, Smeltzer MP. Invasive mediastinal staging for resected non-small cell lung cancer in a population-based cohort. *J Thorac Cardiovasc Surg*. 2019;158(4):1220-1229.e2. https://doi.org/10.1016/j.jtcvs.2019.04.068
13. Zhaoming G, Zhenfa Z. Lymph node evaluation and surgical procedure selection for non-small cell lung cancer. *Holist Integ Oncol*. 2024;3:1-11. https://doi.org/10.1007/s44178-024-00070-3
14. Chen YY, Chen YS, Huang TW. Prognostic impact of EBUS TBNA for lung adenocarcinoma patients with postoperative recurrences. *Diagnostics (Basel)*. 2022;12(10):2547. https://doi.org/10.3390/diagnostics12102547
15. Heiden BT, Eaton DB Jr, Chang SH, et al. Assessment of updated commission on cancer guidelines for intraoperative lymph node sampling in early stage NSCLC. *J Thorac Oncol*. 2022;17(11):1287-1296. https://doi.org/10.1016/j.jtho.2022.08.009
16. Nakajima T, Inage T, Sata Y, et al. Elastography for predicting and localizing nodal metastases during endobronchial ultrasound. *Respiration*. 2015;90(6): 499-506. https://doi.org/10.1159/000441798
17. Riegler J, Labyed Y, Rosenzweig S, et al. Tumor elastography and its association with collagen and the tumor microenvironment. *Clin Cancer Res*. 2018;24(18): 4455-4467. https://doi.org/10.1158/1078-0432.CCR-17-3262
18. Wu J, Sun Y, Wang Y, Ge L, Jin Y, Wang Z. Diagnostic value of endobronchial ultrasound elastography for differentiating benign and malignant hilar and mediastinal lymph nodes: a systematic review and meta-analysis. *Med Ultrason*. 2022; 24(1):85-94. https://doi.org/10.11152/mu-2971
19. Izumo T, Sasada S, Chavez C, Matsumoto Y, Tsuchida T. Endobronchial ultrasound elastography in the diagnosis of mediastinal and hilar lymph nodes. *Jpn J Clin Oncol*. 2014;44(10):956-962. https://doi.org/10.1093/jjco/hyu105
20. Sun J, Zheng X, Mao X, et al. Endobronchial ultrasound elastography for evaluation of intrathoracic lymph nodes: a pilot study. *Respiration*. 2017;93(5): 327-338. https://doi.org/10.1159/000464253
21. Korrungruang P, Boonsarngsuk V. Diagnostic value of endobronchial ultrasound elastography for the differentiation of benign and malignant intrathoracic lymph nodes. *Respirology*. 2017;22(5):972-977. https://doi.org/10.1111/resp.12979
22. Verhoeven RLJ, de Korte CL, van der Heijden EHFM. Optimal endobronchial ultrasound strain elastography assessment strategy: an explorative study. *Respiration*. 2019;97(4):337-347. https://doi.org/10.1159/000494143
23. Mistry N, Gatti A, Churchill I, Patel Y, Hanna W. Determining the optimal stiffness colour threshold and stiffness area ratio cut-off for mediastinal lymph node staging using EBUS elastography and AI: a pilot study. *Can J Surg*. 2021;64(6 Suppl 2):S110.
24. Hylton DA, Turner S, Kidane B, et al; Canadian Association of Thoracic Surgery (CATS) Working Group. The Canada Lymph Node Score for prediction of malignancy in mediastinal lymph nodes during endobronchial ultrasound. *J Thorac Cardiovasc Surg*. 2020;159(6):2499-2507.e3. https://doi.org/10.1016/j.jtcvs.2019.10.205
25. Bujang MA, Adnan TH. Requirements for minimum sample size for sensitivity and specificity analysis. *J Clin Diagn Res*. 2016;10(10):YE01-YE06. https://doi.org/10.7860/JCDR/2016/18129.8744
26. Churchill IF, Gatti AA, Hylton DA, et al. An artificial intelligence algorithm to predict nodal metastasis in lung cancer. *Ann Thorac Surg*. 2022;114(1): 248-256. https://doi.org/10.1016/j.athoracsur.2021.06.082
27. IBM Corp. *IBM SPSS Statistics for Windows, Version 27.0*. IBM Corp; 2020.
28. Mandrekar JN. Receiver operating characteristic curve in diagnostic test assessment. *J Thorac Oncol*. 2010;5(9):1315-1316. https://doi.org/10.1097/JTO.0b013e3181ec173d
29. Šimundić AM. Measures of diagnostic accuracy: basic definitions. *EJIFCC*. 2009;19(4):203-211.
30. Wang Y, Zhao Z, Zhu M, Zhu Q, Yang Z, Chen L. Diagnostic value of endobronchial ultrasound elastography in differentiating between benign and malignant hilar and mediastinal lymph nodes: a retrospective study. *Quant Imaging Med Surg*. 2023;13(7):4648-4662. https://doi.org/10.21037/qims-23-241
31. Uchimura K, Yamasaki K, Sasada S, et al. Quantitative analysis of endobronchial ultrasound elastography in computed tomography-negative mediastinal and hilar lymph nodes. *Thorac Cancer*. 2020;11(9):2590-2599. https://doi.org/10.1111/1759-7714.13579
32. Fujiwara T, Nakajima T, Inage T, et al. The combination of endobronchial elastography and sonographic findings during endobronchial ultrasound-guided transbronchial needle aspiration for predicting nodal metastasis. *Thorac Cancer*. 2019;10(10):2000-2005. https://doi.org/10.1111/1759-7714.13186
33. Ma H, An Z, Xia P, et al. Semi-quantitative analysis of EBUS elastography as a feasible approach in diagnosing mediastinal and hilar lymph nodes of lung cancer patients. *Sci Rep*. 2018;8(1):3571. https://doi.org/10.1038/s41598-018-22006-4
34. Santos MK, Ferreira Júnior JR, Wada DT, et al. Artificial intelligence, machine learning, computer-aided diagnosis, and radiomics: advances in imaging towards to precision medicine. *Radiol Bras*. 2019;52(6):387-396. Erratum in: *Radiol Bras*. 2022;55(3):208. https://doi.org/10.1590/0100-3984.2019.0049
35. Litjens G, Kooi T, Bejnordi BE, et al. A survey on deep learning in medical image analysis. *Med Image Anal*. 2017;42:60-88. https://doi.org/10.1016/j.media.2017.07.005
36. Jha S, Topol EJ. Adapting to artificial intelligence: radiologists and pathologists as information specialists. *JAMA*. 2016;316(22):2353-2354. https://doi.org/10.1001/jama.2016.17438

Key Words: artificial intelligence augmented, clinical utility, endobronchial ultrasound elastography, lung cancer, lymph node staging