Nodal Staging in Lung Cancer

A Risk Stratification Model for Lymph Nodes Classified as Negative by EBUS-TBNA

Matthew Evison,† Julie Morris,‡ Julie Martin,* Rajesh Shah,§ Philip V. Barber,* Richard Booton,*† and Philip A. J. Crosbie*†*

Background: Over the last 10 years, endobronchial ultrasound-guided transbronchial needle aspiration (EBUS-TBNA) has become established as the first-line nodal staging procedure of choice for lung cancer patients. However, the pathway for patients following a negative EBUS-TBNA has not been clearly defined. The primary aim of this study was to develop and validate a risk stratification model to categorize lymph nodes deemed negative by EBUS-TBNA into "low-risk" and "high-risk" groups, where "risk" refers to the risk of false negative sampling.

Methods: A retrospective analysis of a prospectively maintained database at a UK tertiary EBUS-TBNA centre was performed. Only patients with primary lung cancer and only negative lymph nodes by EBUS-TBNA were included in the analysis. A risk stratification model was built from a derivation set using independent predictors of malignancy and the validation set used to evaluate the constructed model. The study period was from March 2010 to August 2013.

Results: Three hundred twenty-nine lymph nodes were included in the analysis (derivation set $n = 196$, validation set $n = 133$). Lymph node standardized uptake value, the standardized uptake value ratio between the lymph node and primary tumor, and heterogeneous echogenicity during sonographic assessment were the only independent predictors of malignancy. Using a simplified scoring system based on the natural logs of the odds ratios from the multivariable analysis on the derivation sample, lymph nodes can be stratified into low risk (score ≤1) and high risk (score ≥2). One hundred forty-one of 142 and 94 of 96 lymph nodes classified as low risk in the derivation and validation set, respectively, were ultimately proven to be benign and 35 of 54 and 24 of 37 lymph nodes classified as high risk were proven malignant. The negative predictive value of the risk stratification model for the derivation set and validation set was 99.3% (95% confidence interval 96.1%–99.6%) and 97.9% (95% confidence interval 92%–99.6%), respectively.

Conclusion: This risk stratification model may assist lung cancer multidisciplinary teams in deciding which patients need further staging procedures and which may proceed directly to treatment after a negative EBUS.

Key Words: Lung cancer, EBUS, EBUS-TBNA, Mediastinal staging, PET-CT.

(*J Thorac Oncol.* 2015;10: 126–133)

Treatment and prognosis in lung cancer is critically dependent on stage. In the absence of distant metastases the staging becomes dependent upon the presence or absence of regional lymph node involvement. Exclusion of N2/3 lymph node metastases allows identification of patients most likely to benefit from surgical resection and hence offer the best chance of long-term cure.[1] Detection of N2/3 nodal metastases identifies those patients that require multimodality therapy.[2]

Endobronchial ultrasound-guided transbronchial needle aspiration (EBUS-TBNA) has led to a paradigm shift in nodal staging pathways in the last 10 years, and is now widely available in the United Kingdom. It has been shown to be safer, cheaper than, and at least as effective as surgical nodal staging when used as the initial investigation of choice.[3–9] Surgical staging, however, is still considered the standard procedure for nodal sampling. Indeed, National guidelines from both the United Kingdom and the United States recommend surgical staging in cases of negative EBUS-TBNA "where the suspicion of malignancy remains high."[10,11] A definition of high suspicion is not provided however. Furthermore, in a randomized control trial of endoscopic staging versus surgical staging, where every negative endoscopic procedure was followed by surgical sampling, 11 patients were required to undergo mediastinoscopy to identify one with N2/3 nodal metastases following a negative endoscopic procedure.[12] Therefore, while the place of EBUS-TBNA as a first-line investigation for invasive nodal staging is well established, the pathway for patients following a negative EBUS-TBNA is not. Considering that

*North West Lung Centre, University Hospital of South Manchester, Manchester, United Kingdom; †The Institute of Inflammation and Repair, The University of Manchester, Manchester, United Kingdom; and Departments of ‡Medical Statistics and §Thoracic Surgery, University Hospital of South Manchester, Manchester, United Kingdom.

Disclosure: Dr. Booton has received grant funding from AstraZeneca and Lilly Oncology and honoraria from Eli-Lilly, AstraZeneca, Chiesi and Almirall. Dr. Evison's fellowship post is half funded by Lilly Oncology. All other authors declare no conflict of interest.

Address for correspondence: Matthew Evison, North West Lung Centre, University Hospital of South Manchester, Southmoor Road, Wythenshawe, Manchester, M23 9LT, United Kingdom. E-mail: matthewevison@hotmail.co.uk, matthew.evison@uhsm.nhs.uk

DOI: 10.1097/JTO.0000000000000348

ISSN: 1556-0864/15/1001-0126

negative EBUS-TBNA represents a significant proportion of all EBUS-TBNA staging procedures (an average of 40% in meta-analysis) and given the majority are truly negative (negative predictive value [NPV] 91%)[11]; if all patients with a negative EBUS-TBNA underwent surgical staging there would a significant number of futile surgical procedures without adding any additional staging information. This could also lengthen the diagnostic pathway and delay definitive treatment. Many authors, therefore, question the validity of surgical staging in all cases of negative EBUS-TBNA.[6,7] Yet the characteristics that may identify which patients would benefit most from further staging have not been defined.

The primary aim of this study was to identify which characteristics may stratify the risk of nodal metastases in lymph nodes deemed negative following EBUS-TBNA. In other words, to provide a definition of "high suspicion" of malignancy that the national guidelines, described above, refer to. If lymph nodes can be accurately stratified into "low-risk" and "high-risk"' categories, where "risk" refers to the risk of false negative sampling, it may help lung cancer multidisciplinary teams (MDTs) to define which patients may need further surgical staging after EBUS-TBNA and those that may proceed directly to definitive treatment without further staging procedures.

PATIENTS AND METHODS

The University Hospital South Manchester (UHSM) provides a tertiary EBUS-TBNA service for a large Cancer Network in the North West of England. Prospective data are collected on all patients undergoing EBUS-TBNA and has been since its launch in March 2010. Patient demographics, referral and pathway data, and radiological characteristics are recorded before the procedure. Intraoperative findings, nodal sampling data, and complications are recorded following completion of the procedure and patient recovery. The pathological results of EBUS-TBNA, the results of any subsequent nodal sampling, e.g., mediastinoscopy and the outcome of 6 months of clinical–radiological follow-up are then added to the database when available. This study is a retrospective analysis of this prospectively maintained database.

Patient Selection

All patients referred to UHSM for EBUS-TBNA for nodal staging of suspected or confirmed lung cancer were reviewed. However, only patients with a pathological diagnosis of primary lung cancer or an MDT-agreed clinical diagnosis of lung cancer, such that definitive treatment was recommended, were included in the analysis. The reason for this was to simulate how we envisage a risk stratification process being used in real life practice. The appropriate time to use this stratification would be when all minimally invasive staging results, both radiological and pathological, are available, most likely in an MDT setting. This prevents patients with benign disease or a low suspicion of lung cancer undergoing nodal staging stratification, when the focus of care may be elsewhere, such as further diagnostics or radiological follow-up. Patients diagnosed with small-cell lung cancer were kept in the analysis as risk stratification in cases of negative EBUS-TBNA is relevant to this cohort of patients, as well as non–small-cell lung cancer (NSCLC) patients.

Nodal Characteristics

For every patient undergoing EBUS-TBNA, the following information is collected for each lymph node sampled: size in short axis on computed tomography (CT), maximum standardized uptake value (SUVmax) of the primary tumor and SUVmax of the lymph node (in those patients undergoing positron emission tomography-computed tomography [PET-CT] as part of their staging pathway) and lymph node appearances using endobronchial ultrasound. In 2010, Fujiwara et al.[13] published the first analysis of lymph node appearances using endobronchial ultrasound. This large, retrospective study investigated five potential sonographic characteristics (Table 1), four of which were shown to be independent predictors of nodal metastases.[13] The same five characteristics have been collected since the launch of EBUS-TBNA at UHSM; echogenicity, shape, margin, central hilar structure, and coagulation necrosis sign.

Endobronchial Ultrasound-Guided Transbronchial Needle Aspiration

The accepted indications for EBUS-TBNA nodal staging at UHSM are: any hilar or mediastinal node with a short axis greater than 10 mm on CT, any abnormal fluorodeoxyglucose (FDG) activity in a hilar or mediastinal lymph node above that of the mediastinal blood pool and a central tumor with a normal mediastinum (defined as within the inner third of the thorax on transverse CT). Peripheral tumors with a normal mediastinum (defined as within the outer two thirds of the thorax on transverse CT) do not undergo pathological staging

TABLE 1. Definitions of the Sonographic Characteristics Used in the Assessment of Lymph Nodes During EBUS-TBNA

Sonographic Feature	Description	Predictor of Malignancy
Echogenicity	Homogeneous or heterogeneous appearance to the lymph node	Heterogeneous
Shape (oval vs. round)	Ratio of the short axis to the long axis of the lymph node	Round shape (ratio <1.5)
Margins (distinct vs. indistinct)	An echogenic border is visible between the lymph node and adjacent tissue	Distinct margin (>50% of the echoic border is visible)
Central hilar structure	A linear, hyperechoic area in the centre of the lymph node	Absence of the central hilar structure
Coagulation necrosis sign	A hypoechoic area without blood flow (representing necrosis), usually in the periphery of the lymph node	Presence of the coagulation necrosis sign

EBUS-TBNA, endobronchial ultrasound-guided transbronchial needle aspiration.

and proceed straight to surgical resection assuming adequate fitness and the absence of distant metastases. EBUS-TBNA is performed under conscious sedation, using incremental doses of midazolam and alfentanyl, without anesthetic support. Standard diagnostic bronchoscopy is performed before all EBUS procedures using Olympus (BFF260 or BF6C260; Olympus, Southend-on-sea, Essex, United Kingdom) bronchoscopes. EBUS-TBNA is performed using an Olympus BF-UC260FW ultrasonic bronchoscope and either 21- or 22-gauge needles. A systematic examination of the mediastinum is performed and any lymph nodes that meet any of the following criteria are sampled: greater than 10 mm short axis on CT or during "real-time" ultrasound assessment, abnormal nodal FDG avidity above that of the mediastinal blood pool on PET or any abnormal ultrasound appearances according to the image classification system proposed by Fujiwara et al.[13] N3 nodes are sampled before N2 nodes followed by N1 nodes. Sampling of lymph nodes from the esophagus is not performed and rapid onsite evaluation of samples is not available.

Statistical Analysis

The following characteristics were analyzed as potential predictors in a risk stratification model: lymph node size (mm in short axis on CT), lymph node SUV, SUV ratio (ratio between the SUV of the lymph node and the primary tumor, e.g., if the lymph node has a SUV of 3.0 and the primary tumor 12.0 the ratio is 0.25, 25%), and sonographic characteristics (echogenicity, shape, margin, central hilar structure, and coagulation necrosis sign). Individual lymph nodes were classified as "benign" or "malignant" based on the pathological results of EBUS-TBNA and any further pathological sampling plus 6 months of clinical–radiological follow-up. A lymph node was only considered benign if all pathological results and 6 months of follow-up failed to yield any evidence of malignancy. The statistical analysis consisted of two stages. First, each individual characteristic was first assessed for its staging performance on the entire cohort of lymph nodes using standard definitions. To develop the risk stratification model only lymph nodes that were negative or inadequately sampled by EBUS-TBNA were used in the subsequent analysis. The decision to incorporate inadequate samples into the risk stratification model allows all patients with no evidence of malignancy from EBUS-TBNA to be incorporated. Furthermore, the classification of an EBUS-TBNA sample as inadequate is a subjective opinion by the reporting pathologist and objective definitions for inadequate sampling are lacking. All negative or inadequate lymph nodes were randomly divided according to a 60%:40% ratio into a derivation set and a validation set. Logistic regression analysis was then applied to the derivation set to identify significant independent predictors of malignancy and then construct a prognostic index using the coefficients from the regression analysis which would give an optimal prediction score for malignancy. An appropriate score cutoff was then assessed. The validation sample was then used to evaluate the diagnostic accuracy of the constructed score.

In addition to the analysis described above, an interobserver agreement for the interpretation of ultrasound characteristics was performed. Two EBUS operators, blinded to the final lymph node pathology, were asked to interpret the ultrasound images recorded during EBUS for 50 lymph nodes. These were randomly selected by a computer program from the 329 negative or inadequate lymph nodes used in the risk stratification modeling. The data were summarized using frequencies, percentages and cross tabulations, and the percentage of observed overall agreement. In addition, the agreement was analyzed using Cohen's Kappa statistic. The analyses used the conventional two-sided 5% significance level. All summaries and analyses were produced using SPSS version 20.

Sample Size

A formal sample size was not calculated but the power of the analysis is demonstrated by the CIs of the performance indicators (calculated using the Wilson method).

Study Period

The study period chosen was from March 2010 to August 2013. The analysis was undertaken in March 2014 to allow 6 months of clinical–radiological follow-up for all lymph nodes.

RESULTS

Patient Characteristics

A total of 509 patients with lung cancer underwent EBUS-TBNA in the study period. All 509 patients underwent CT thorax before EBUS-TBNA and 72% (365 of 509) also underwent PET-CT in their staging work-up. The majority of patients had discrete enlargement of mediastinal lymph nodes on CT (74%, 376 of 509, CT N2/3) and abnormal FDG avidity within mediastinal lymph nodes (86%, 313 of 365 PET-CT N2/3). Patient characteristics are presented in Table 2. The mean age was 70 years and 51% (261/509) were male. The majority were of good performance status (68%, 347 of 509 performance status 0–1). Ultimately, 78% (396 of 509) were pathologically proven to have NSCLC, 5% (28 of 509) were not pathologically proven but treated under a clinical diagnosis of NSCLC and 17% were diagnosed with small-cell lung cancer. A total of 877 lymph nodes were sampled (1.7 per patient). The mean size of lymph node sampled was 17 ± 8 mm. Sixty-six percent (337 of 509) of patients were ultimately staged with advanced nodal disease (N2/3) and 68% of lymph nodes sampled were ultimately proven to be malignant (596 of 877).

CT Nodal Staging

Of the 877 lymph nodes sampled with EBUS-TBNA, 22% (190 of 877) were less than or equal to 10 mm and 78% (687 of 877) were greater than 10 mm. The mean size for benign lymph nodes was 11 ± 3 mm compared with a mean size of 19 ± 9 mm for malignant lymph nodes. One hundred thirty-seven of 190 lymph nodes measuring less than or equal to 10 mm on CT were ultimately classified as benign. Five hundred forty-three of 687 lymph nodes measuring greater than 10 mm on CT were ultimately classified as malignant. A lymph size greater than 10 mm in short axis on CT was associated with an increased risk of malignancy (OR 9.7, 95% confidence interval [CI] 6.8–14.1). The performance

TABLE 2. Patient Characteristics

Age (mean ± SD)	69.9 ± 9.6
Sex, *n* (%)	
Male	261 (51.3)
Female	248 (48.7)
Performance status (*n* = 509), *n* (%)	
0	84 (16.5)
1	263 (51.7)
2	131 (25.7)
3	29 (5.7)
4	2 (0.4)
Histological subtyping (*n* = 509), *n* (%)	
Adenocarcinoma	195 (38.3)
Squamous cell carcinoma	165 (32.4)
NSCLC "not otherwise specified"	27 (5.3)
Large cell carcinoma	9 (1.8)
Clinical diagnosis of NSCLC (without pathological confirmation)	28 (5.5)
Small-cell lung cancer	85 (16.7)
Final nodal staging (*n* = 509), *n* (%)	
N0	112 (22)
N1	60 (11.8)
N2	246 (48.3)
N3	91 (17.9)
Lymph node stations sampled with EBUS-TBNA (*n* = 877), *n* (%)	
High right paratracheal (2R)	29 (3.3)
High left paratracheal (2L)	1 (0.1)
Retrotracheal (3p)	8 (0.9)
Low right paratracheal (4R)	256 (29.2)
Low left paratracheal (4L)	123 (14.0)
Subcarinal (7)	285 (32.5)
Right hilar (10R, 11R)	106 (12.1)
Left hilar (10L, 11L)	69 (7.9)

EBUS-TBNA, endobronchial ultrasound-guided transbronchial needle aspiration; NSCLC, non–small-cell lung cancer.

characteristics of CT nodal staging were as follows: sensitivity 91.1%, specificity 48.8%, NPV 72.1%, PPV 79.0%, and diagnostic accuracy 77.5%. The receiver operating characteristic (ROC) curve demonstrates an area under the curve of 0.843 (95% CI 0.816–0.869, Figure 1).

PET Nodal Staging–Lymph Node SUV

A total of 365 patients underwent PET-CT as part of their staging work-up, and 675 lymph nodes were sampled from these patients. The mean lymph node SUV was 7.3 ± 5.8. In those lymph nodes ultimately proven to be benign the mean SUV was 2.9 ± 1.2 compared with 10.1 ± 5.9 in malignant lymph nodes. A total of 564 lymph nodes were "PET positive," defined as an SUV above that of the mediastinal blood pool, and 111 lymph nodes "PET negative." Three hundred ninety-seven of 564 PET-positive lymph nodes were ultimately proven to be malignant and 102 of 111 PET-negative lymph nodes proven benign. A lymph node SUV greater than that of the mediastinal blood pool was associated with an increased probability of malignancy (OR 26.9, 95% CI 13.3–54.5). The diagnostic performance of PET-CT nodal staging in this study was as follows: sensitivity 97.8%, specificity 37.9%, NPV 91.9%, PPV 70.4%, and diagnostic accuracy 73.9%. The ROC curve demonstrates an area under the curve of 0.944 (95% CI 0.927–0.961, Figure 1). Following interrogation of the ROC curve a different "cutoff" value to indicate "PET positivity" was analyzed. Using an SUV value greater than 4.0 to indicate PET positivity the diagnostic performance was as follows: sensitivity 89.9%, specificity 89.6, NPV 85.8, PPV 92.6%, diagnostic accuracy 89.8%, and OR 74.0 (95% CI 44.8–122.3).

PET Nodal Staging–SUV Ratio

For 673 lymph nodes it was possible to calculate the SUV ratio between the lymph node and the primary tumor, expressed as a percentage. The mean ratio was 65.8% ± 53.8%. In those lymph nodes ultimately proven to be benign the mean ratio was 26.7% ± 21.8% compared with 91% ± 53.0% in malignant lymph nodes. The ROC curve demonstrates an area under the curve of 0.928 (95% CI 0.907–0.950, Figure 1). Following interrogation of the ROC curve an optimal cutoff point was analyzed to indicate malignancy. Using an SUV ratio of greater than 40% to indicate malignancy the diagnostic performance was as follows: sensitivity 90.9%, specificity 86.6%, NPV 86.2%, PPV 91.1%, diagnostic accuracy 89.2%, and OR 64.1 (95% CI 34.1–104.4).

Endobronchial Sonographic Staging

A complete data set for ultrasound characteristics was achieved in 858 lymph nodes (98%, 858 of 877) and used in the analysis. The prevalence of benign and malignant lymph nodes stratified according each individual sonographic characteristic is presented in Figure 1 and the diagnostic performance of each characteristic in Table 3. Heterogeneous echogenicity (OR 123.0, 95% CI 64.8–233.6), round shape (OR 11.7, 95% CI 8.1–16.8), distinct margin (OR 1.4, 95% CI 1.1–1.9), the absence of a central hilar structure (OR 18.5, 95% CI 10.8–31.7), and the presence of a coagulation necrosis sign (OR 50.4, 95% CI 7.0–363.7) were all associated with an increased risk of malignancy.

EBUS-TBNA Nodal Staging

A total of 877 lymph nodes were sampled with EBUS-TBNA. Five percent (46 of 877) were deemed inadequate by the reporting pathologist. Nodal metastases were confirmed by EBUS-TBNA in 60% (525 of 877). Thirty-five percent (306 of 877) of lymph nodes were classified as negative by EBUS-TBNA and 16% (48 of 306) were subsequently proven to be false negatives. The performance characteristics of EBUS-TBNA for nodal staging in this study were as follows: sensitivity 91.6%, NPV 84.3%, and diagnostic accuracy 94.2% (specificity and PPV are assumed to be 100%). The verification method for the final lymph node pathological diagnosis in cases of negative or inadequate lymph node sampling was

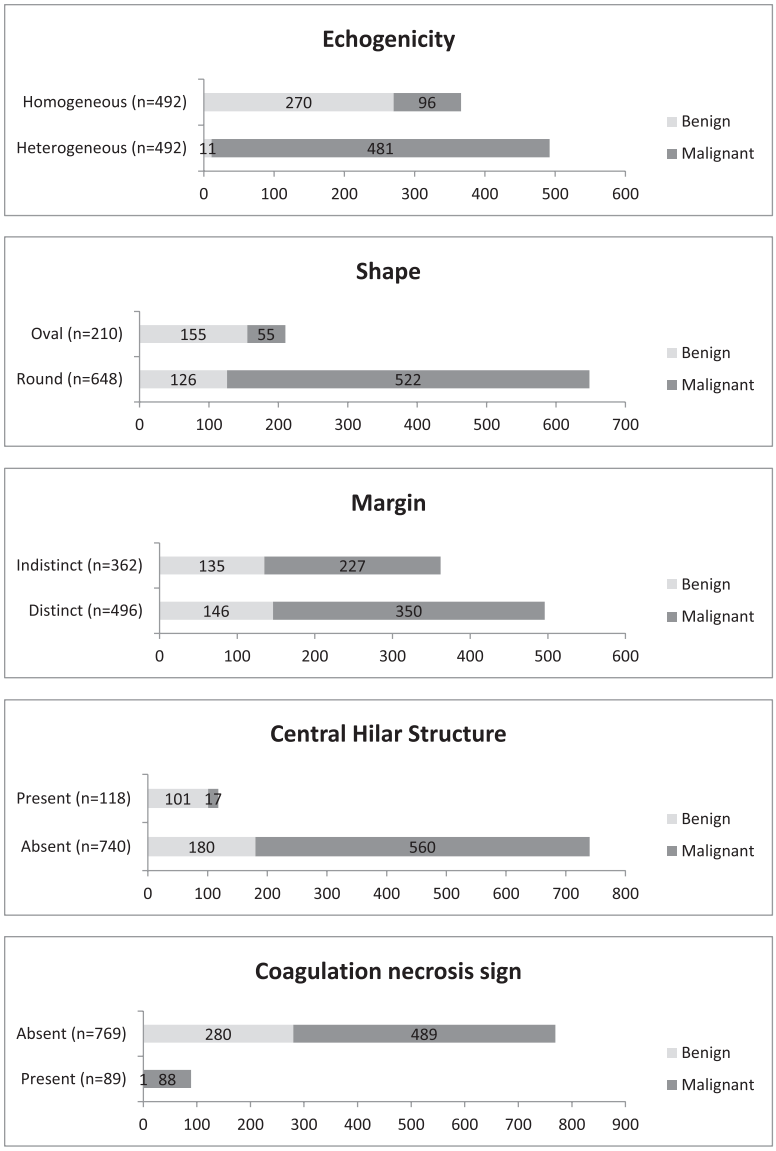

FIGURE 1. Ultrasound characteristics stratified by lymph node pathology, benign versus malignant.

surgical sampling in 49% (174 of 352) and clinical–radiological in 51% (178 of 352).

Risk Stratification Model in Negative or Inadequate Lymph Nodes

From the total sample of 877 lymph nodes, 352 were classified as negative or inadequate by EBUS-TBNA. Of these, 329 lymph nodes had complete data for every potential predictive characteristic. These 329 lymph nodes were randomly divided according to a 60:40 ratio into a derivation set (n = 196) and a validation set (n = 133). The derivation set contained 36 malignant lymph nodes (18%), and the validation set contained 26 malignant lymph nodes (19%). The logistic regression analysis performed on the derivation set is presented in Table 4. The coagulation necrosis sign was not included in the analysis because there were too few lymph nodes with this feature present to enable a reasonably robust regression analysis with this factor. Lymph node SUV, SUV ratio, and echogenicity were independent predictors of nodal metastases. Using a simplified scoring system (Table 5) based on the natural logs of the odds ratios from the previous multivariable analysis on the derivation sample, lymph nodes can be stratified into low risk (score ≤1) and high risk (score ≥2). In the derivation set, 142 lymph nodes were classified as low risk and 99.3% (141 of 142) were ultimately proven to be benign (NPV 99.3%, 95% CI 96.1–99.6). Fifty-four lymph nodes were classified as high risk and 64.8% (35 of 54) were ultimately proven to be malignant (positive predictive value [PPV] 64.8%, 95% CI 52%–76%). The sensitivity of this scoring system on the derivation set was therefore 97.2% (35 of 36, 95% CI 86%–100%) and the specificity 88.1% (141 of 160, 95% CI 82%–92%).

The same scoring system was applied to the validation set of 133 lymph nodes. Ninety-six lymph nodes were classified as low risk and 97.9% (94 of 96) were ultimately proven to be benign (NPV 97.9%, 95% CI 92%–99.6%).

TABLE 3. Diagnostic Performance for Each Ultrasound Characteristic

Characteristic	Sensitivity (%)	Specificity (%)	NPV (%)	PPV (%)	Accuracy (%)
Echogenicity	83.4	96.1	73.8	97.8	87.5
Shape	90.5	55.2	73.8	80.6	78.9
Margin	60.7	48.0	37.3	70.6	56.5
CHS	97.1	35.9	88.6	75.7	77.0
CNS	15.3	99.6	36.4	98.9	42.9

NPV, negative predictive value; PPV, positive predictive value; CHS, central hilar structure; CNS, coagulation necrosis sign.

Thirty-seven lymph nodes were classified as high risk and 64.9% (24 of 37) were ultimately proven to malignant (PPV 64.9%, 95% CI 49%–78%). The sensitivity of the scoring system in the validation set was 92.3% (24 of 26, 95% CI 76%–98%) and the specificity 87.8% (94 of 107, 95% CI 80%–93%).

TABLE 4. Logistic Regression Analysis on Derivation Sample (n = 196) to Predict Malignancy

	Univariate Analysis		Multivariable Analysis	
	OR (95% CI)	Significance	OR[a] (95% CI)	Significance
Echogenicity				
Homogeneous (ref)	1	$p < 0.001$	1	$p < 0.001$
Heterogeneous	135 (40–454)		48 (8–282)	
SUV				
≤4 (ref)	1	$p < 0.001$	1	$p = 0.007$
>4	60 (20–177)		10 (1.9–59)	
Lymph SUV%				
≤40 (ref)	1	$p < 0.001$	1	$p = 0.001$
41–60	32 (9–115)		9 (1.2–71)	
>60	81 (23–287)		46 (5–379)	
Size				
≤10 mm (ref)	1	$p = 0.002$		
>10 mm	4.0 (1.7–9.8)			
Shape				
Oval (ref)	1	$p < 0.001$		
Round	21 (5–92)			
Margin				
Distinct (ref)	1	$p = 0.57$		
Indistinct	1.2 (0.6–2.6)			
CHS				
Present (ref)	1	$p = 0.005$		
Absent	18 (2–137)			

[a]Adjusting for all other factors in the multivariable model.
CI, confidence interval; SUV, standardized uptake value.

TABLE 5. Lymph Node Scoring System for Predicting Malignancy in Patients Not Found Malignant by EBUS-TBNA

	0	1	2
Echogenicity	Homogeneous		Heterogeneous
SUV	≤4	>4	
Lymph SUV%	≤40	41–60	>60

Total score varies between 0 and 5.
EBUS-TBNA, endobronchial ultrasound-guided transbronchial needle aspiration; SUV, standardized uptake value.

Interobserver Agreement for Ultrasound Characteristics

The strength of agreement between operators was "very good" for all characteristics except "margin" (Table 6). In particular, the observed agreement for echogenicity, the only ultrasound characteristic to be an independent predictor of nodal malignancy in multivariate analysis, was 93% (kappa 0.86, 95% CI 0.7–1.0, $p < 0.001$).

DISCUSSION

It has been well documented that lymph size on CT is a poor predictor of nodal metastases in lung cancer.[11,14] It is therefore unsurprising to find similar results in this study. The limitations of PET-CT in nodal staging are also well documented and consistent with our results. However, previous studies have investigated alternative methods of PET-CT interpretation, just as this study has. Bryant et al.[15] prospectively examined 397 patients under assessment for surgical resection of primary lung cancer in a single centre. The SUVmax of lymph nodes was evaluated against the pathological diagnosis and using a lymph node SUVmax cutoff of greater than or equal to 5.3 to predict nodal metastases achieved a sensitivity of 91% and specificity of 88%.[15] The SUV ratio was first investigated by Cerfolio et al. in patients with NSCLC and an SUVmax within at least one mediastinal lymph node of greater than 2.5. The prevalence of nodal malignancy in this study was 78%. The mean SUV ratio in malignant lymph nodes was 58% compared with 40% in benign nodes. The authors reported an area under the curve of 0.79, a sensitivity of 94%, and a specificity of 72% when a ratio cutoff

TABLE 6. Analysis of Interobserver Agreement for Ultrasound Characteristics

	Observed Agreement	Kappa	95% CI	p Value	Strength of Agreement
Echogenicity	93%	0.86	0.7–1.0	<0.001	Very good
Shape	93%	0.85	0.69–1.0	<0.001	Very good
Margin	69.8%	0.4	0.13–0.67	0.007	Fair
CHS	95.3%	0.89	0.74–1.0	<0.001	Very good
CNS	95.3%	0.83	0.6–1.0	<0.001	Very good

CHS, central hilar structure; CNS, coagulation necrosis sign.

for malignancy of greater than or equal to 56% was used.[16] Koksal et al.[17] subsequently undertook a retrospective review of 100 lymph nodes from NSCLC patients that had undergone surgical resection and lymph node dissection. The prevalence of nodal malignancy was low at 14% (14 of 100). The authors reported an area under the curve of 0.72 for the SUV ratio and if a cutoff of 20% was used a sensitivity of 93% and a specificity of 47% was achieved.[17] In our study, using an SUV cutoff of greater than 4.0 and SUV ratio greater than 40% the diagnostic performance was excellent and improved the specificity of PET-CT from 38% to approximately 90% while maintaining good sensitivity (86%–89%). The prevalence of malignancy in our study was 68% and the results are similar to those reported by Cerfolio, in which the prevalence of nodal malignancy was similarly high.

In our study, echogenicity was the strongest sonographic predictor of nodal malignancy from Fujiwara's proposed EBUS image classification system. It was the only sonographic characteristic that proved to be an independent predictor of malignancy on multivariate analysis. Further studies on sonographic characteristics been published after this study began, the results of which are conflicting. Wang Memoli et al. prospectively evaluated the sonographic characteristics in 227 lymph nodes from 100 lung cancer patients undergoing EBUS staging.[18] Only lymph node shape was found to be predictive of nodal metastases, with no predictive power from margin or echogenicity. There was no interobserver analysis for this study however. A further prospective study by Schmid-Bindert et al.[19] investigated shape, margin, echogenicity, central hilar structure, and the color power Doppler index in 281 lymph nodes from 145 patients. The strength of this study was the surgical verification of lymph node pathology in 148 of 162 lymph nodes that were negative by EBUS-TBNA and a robust analysis of interobserver interpretation. As in our study, echogenicity was the best performer with 85% of heterogeneous lymph nodes proven to be malignant. The color power Doppler index showed both poor predictive power for nodal metastases and only moderate agreement between operators. In contrast, Nakajima et al.[20] reported more promising results using the power Doppler mode in a retrospective study of 100 patients and, most recently, the first preliminary report for endobronchial ultrasound elastography, described as computer-assisted palpation, has been published.[21]

Regardless of these findings, this study was not designed to investigate the utility of radiological methods for nodal staging, through CT, PET-CT, or ultrasound. Radiological staging will never replace pathological staging. However, in cases of negative or inadequate EBUS-TBNA sampling, this study demonstrates the combination of PET-CT data and ultrasound data can stratify patients into high risk and low risk for nodal malignancy. Lymph nodes stratified as low risk by this stratification model have between a 98% and 99% chance of being truly negative based on our validation and derivation sets, respectively. The lowest value in the confidence interval of either calculation was 92%, higher than the NPV of EBUS-TBNA alone (84%). The PPV is approximately 65% in both cohorts. This is due to the cutoff score chosen to separate high risk and low risk and is a necessity to maintain such a high NPV. It is surely better to undertake mediastinoscopy and find benign lymph nodes than proceed with inappropriate radical treatment due to undetected malignant nodes. There are a number of discussion points when considering these results. First, this analysis is based on a per lymph node basis whereas much of the EBUS-TBNA literature is reported on a per patient basis. The denominator for per patient outcomes is the presence or absence of N2/3 nodal metastases. However, at our centre, single station or single zone N2 is managed differently form multistation N2, with surgical resection followed by adjuvant chemotherapy preferred to concurrent chemoradiotherapy. We believe, therefore, EBUS nodal staging must accurately diagnose the nodal metastases within the single station N2 nodes and exclude metastases at other N2/3 nodes. To reflect the success or failure of this requires a per lymph analysis. In addition, there will frequently be differing SUVs and ultrasound characteristics within the same patient, again necessitating each individual lymph node to be considered independently for risk stratification. Second, the verification method in our study is suboptimal with only 49% of negative lymph nodes verified surgically. In an ideal world, all nodes would have been surgically verified but in the real world this is not possible. Many patients will not be fit for surgical procedures, indeed 32% of patients were PS of greater than or equal to 2, and many go on to have oncological management that lacks any surgical involvement. Finally, but most importantly, there are limitations to consider in terms of which lymph nodes were sampled during EBUS procedures and the quality of the staging technique undertaken with EBUS. This study only assessed lymph nodes sampled with EBUS-TBNA, not all lymph nodes. Therefore, both lymph node stations inaccessible by EBUS-TBNA and lymph nodes examined during EBUS but not sampled are not included in the analysis. Given the criteria, set out at the beginning of this study, the lymph nodes that were examined during EBUS but not sampled were less than 10 mm, PET negative with no abnormal sonographic features. This resulted in an average of 1.7 lymph nodes sampled per patient. However, our practice has changed towards the later stages of this study, particularly in light of the revised European Society of Thoracic Surgeons Guidelines on preoperative nodal staging in lung cancer.[22] These guidelines mandate the sampling of any lymph nodes measuring greater than 5 mm during endoscopic assessment, with a minimum of three lymph node stations sampled (4R, 7, and 4L). The majority of lymph nodes in this study were enlarged or FDG avid. The risk stratification model, therefore, can only be applied to such lymph nodes and has not been proven in small, PET-negative nodes. The authors are committed to ongoing data collection and analysis of this risk stratification model in the context of our revised practice and systematic nodal sampling. We intend to interrogate the effectiveness of this model in small, PET-negative nodes with further publications planned.

Ultimately, nodal staging pathways will be heavily dependent on local expertise and services. This risk stratification model provides a mechanism for lung cancer MDTs to discuss the risk of false negative EBUS-TBNA sampling in cases of enlarged or PET positive lymph nodes but negative or inadequate EBUS-TBNA pathology. This may aid the

decision making with regard to further staging procedures. It is dependent on EBUS operators being proficient in the interpretation of sonographic characteristics during EBUS and robust reporting of these findings.

REFERENCES

1. Howington JA, Blum MG, Chang AC, Balekian AA, Murthy SC. Treatment of stage I and II 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):e278S–e313S.
2. Ramnath N, Dilling TJ, Harris LJ, et al. Treatment of stage III 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):e314S–e340S.
3. Adams K, Shah PL, Edmonds L, Lim E. Test performance of endobronchial ultrasound and transbronchial needle aspiration biopsy for mediastinal staging in patients with lung cancer: systematic review and meta-analysis. *Thorax* 2009;64:757–762.
4. Gu P, Zhao YZ, Jiang LY, Zhang W, Xin Y, Han BH. Endobronchial ultrasound-guided transbronchial needle aspiration for staging of lung cancer: a systematic review and meta-analysis. *Eur J Cancer* 2009;45:1389–1396.
5. Varela-Lema L, Fernández-Villar A, Ruano-Ravina A. Effectiveness and safety of endobronchial ultrasound-transbronchial needle aspiration: a systematic review. *Eur Respir J* 2009;33:1156–1164.
6. Yasufuku K, Pierre A, Darling G, et al. A prospective controlled trial of endobronchial ultrasound-guided transbronchial needle aspiration compared with mediastinoscopy for mediastinal lymph node staging of lung cancer. *J Thorac Cardiovasc Surg* 2011;142:1393–400.e1.
7. Ernst A, Anantham D, Eberhardt R, Krasnik M, Herth FJ. Diagnosis of mediastinal adenopathy-real-time endobronchial ultrasound guided needle aspiration versus mediastinoscopy. *J Thorac Oncol* 2008;3:577–582.
8. Harewood GC, Pascual J, Raimondo M, et al. Economic analysis of combined endoscopic and endobronchial ultrasound in the evaluation of patients with suspected non-small cell lung cancer. *Lung Cancer* 2010;67:366–371.
9. Steinfort DP, Liew D, Conron M, Hutchinson AF, Irving LB. Cost-benefit of minimally invasive staging of non-small cell lung cancer: a decision tree sensitivity analysis. *J Thorac Oncol* 2010;5:1564–1570.
10. NICE Clinical Guidelines. *The Diagnosis and Treatment of Lung Cancer (Update)*. Cardiff, UK: National Collaborating Centre for Cancer, 2011.
11. 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.
12. Annema JT, van Meerbeeck JP, Rintoul RC, et al. Mediastinoscopy vs endosonography for mediastinal nodal staging of lung cancer: a randomized trial. *JAMA* 2010;304:2245–2252.
13. Fujiwara T, Yasufuku K, Nakajima T, et al. The utility of sonographic features during endobronchial ultrasound-guided transbronchial needle aspiration for lymph node staging in patients with lung cancer: a standard endobronchial ultrasound image classification system. *Chest* 2010;138:641–647.
14. Navani N, Spiro SG, Janes SM. Mediastinal staging of NSCLC with endoscopic and endobronchial ultrasound. *Nat Rev Clin Oncol* 2009;6:278–286.
15. Bryant AS, Cerfolio RJ, Klemm KM, Ojha B. Maximum standard uptake value of mediastinal lymph nodes on integrated FDG-PET-CT predicts pathology in patients with non-small cell lung cancer. *Ann Thorac Surg* 2006;82:417–422; discussion 422–423.
16. Cerfolio RJ, Bryant AS. Ratio of the maximum standardized uptake value on FDG-PET of the mediastinal (N2) lymph nodes to the primary tumor may be a universal predictor of nodal malignancy in patients with nonsmall-cell lung cancer. *Ann Thorac Surg* 2007;83:1826–1829; discussion 1829–1830.
17. Koksal D, Demirag F, Bayiz H, et al. The correlation of SUVmax with pathological characteristics of primary tumor and the value of tumor/lymph node SUVmax ratio for predicting metastasis to lymph nodes in resected NSCLC patients. *J Cardiothorac Surg* 2013;8:63.
18. Memoli JS, El-Bayoumi E, Pastis NJ, et al. Using endobronchial ultrasound features to predict lymph node metastasis in patients with lung cancer. *Chest* 2011;140:1550–1556.
19. Schmid-Bindert G, Jiang H, Kähler G, et al. Predicting malignancy in mediastinal lymph nodes by endobronchial ultrasound: a new ultrasound scoring system. *Respirology* 2012;17:1190–1198.
20. Nakajima T, Anayama T, Shingyoji M, Kimura H, Yoshino I, Yasufuku K. Vascular image patterns of lymph nodes for the prediction of metastatic disease during EBUS-TBNA for mediastinal staging of lung cancer. *J Thorac Oncol* 2012;7:1009–1014.
21. Trosini-Désert V, Jeny F, Taillade L, et al. Bronchial endoscopic ultrasound elastography: preliminary feasibility data. *Eur Respir J* 2013;41:477–479.
22. 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;45:787–798.

