

# A Prediction Model to Help with Oncologic Mediastinal Evaluation for Radiation: HOMER

Gabriela Martinez-Zayas<sup>1,2</sup>, Francisco A. Almeida<sup>3</sup>, Michael J. Simoff<sup>4</sup>, Lonny Yarmus<sup>5</sup>, Sofia Molina<sup>1,2</sup>, Benjamin Young<sup>6</sup>, David Feller-Kopman<sup>5</sup>, Ala-Eddin S. Sagar<sup>2</sup>, Thomas Gildea<sup>3</sup>, Labib G. Debiase<sup>5</sup>, Horiana B. Grosu<sup>2</sup>, Roberto F. Casal<sup>2</sup>, Muhammad H. Arain<sup>2</sup>, George A. Eapen<sup>2</sup>, Carlos A. Jimenez<sup>2</sup>, Laila Z. Noor<sup>2</sup>, Shiva Baghaie<sup>2</sup>, Juhee Song<sup>7</sup>, Liang Li<sup>7</sup>, and David E. Ost<sup>2</sup>

<sup>1</sup>Escuela de Medicina y Ciencias de la Salud, Tecnológico de Monterrey, Monterrey, Mexico; <sup>2</sup>Department of Pulmonary Medicine and <sup>7</sup>Department of Biostatistics, The University of Texas MD Anderson Cancer Center, Houston, Texas; <sup>3</sup>Department of Pulmonary Medicine, Cleveland Clinic, Cleveland, Ohio; <sup>4</sup>Department of Pulmonary and Critical Care Medicine, Henry Ford Hospital, Detroit, Michigan; <sup>5</sup>Division of Pulmonary and Critical Care, Johns Hopkins University, Baltimore, Maryland; and <sup>6</sup>Division of Pulmonary and Critical Care Medicine, University Hospitals, Cleveland, Ohio

ORCID ID: 0000-0001-6351-5552 (D.E.O.).

## Abstract

**Rationale:** When stereotactic ablative radiotherapy is an option for patients with non-small cell lung cancer (NSCLC), distinguishing between N0, N1, and N2 or N3 (N2|3) disease is important.

**Objectives:** To develop a prediction model for estimating the probability of N0, N1, and N2|3 disease.

**Methods:** Consecutive patients with clinical-radiographic stage T1 to T3, N0 to N3, and M0 NSCLC who underwent endobronchial ultrasound-guided staging from a single center were included. Multivariate ordinal logistic regression analysis was used to predict the presence of N0, N1, or N2|3 disease. Temporal validation used consecutive patients from 3 years later at the same center. External validation used three other hospitals.

**Measurements and Main Results:** In the model development cohort ( $n = 633$ ), younger age, central location, adenocarcinoma, and higher positron emission tomography-computed tomography nodal stage were

associated with a higher probability of having advanced nodal disease. Areas under the receiver operating characteristic curve (AUCs) were 0.84 and 0.86 for predicting N1 or higher (vs. N0) disease and N2|3 (vs. N0 or N1) disease, respectively. Model fit was acceptable (Hosmer-Lemeshow,  $P = 0.960$ ; Brier score, 0.36). In the temporal validation cohort ( $n = 473$ ), AUCs were 0.86 and 0.88. Model fit was acceptable (Hosmer-Lemeshow,  $P = 0.172$ ; Brier score, 0.30). In the external validation cohort ( $n = 722$ ), AUCs were 0.86 and 0.88 but required calibration (Hosmer-Lemeshow,  $P < 0.001$ ; Brier score, 0.38). Calibration using the general calibration method resulted in acceptable model fit (Hosmer-Lemeshow,  $P = 0.094$ ; Brier score, 0.34).

**Conclusions:** This prediction model can estimate the probability of N0, N1, and N2|3 disease in patients with NSCLC. The model has the potential to facilitate decision-making in patients with NSCLC when stereotactic ablative radiotherapy is an option.

**Keywords:** lung cancer; lung cancer staging; endobronchial ultrasound; mediastinal adenopathy

In patients with non-small cell lung cancer (NSCLC), correct staging is necessary to offer appropriate treatment. Treatment options and prognosis vary significantly by

stage (1). After ruling out metastatic disease, knowing the N stage is necessary to determine the best treatment strategy. Patients with stage I or a subset of stage II

(T1–2, N1) NSCLC are generally candidates for surgical resection and mediastinal lymph node dissection. For patients with early-stage NSCLC who are medically

(Received in original form April 17, 2019; accepted in final form September 27, 2019)

Ⓐ This article is open access and distributed under the terms of the Creative Commons Attribution Non-Commercial No Derivatives License 4.0 (<http://creativecommons.org/licenses/by-nc-nd/4.0/>). For commercial usage and reprints, please contact Diane Gern ([dgern@thoracic.org](mailto:dgern@thoracic.org)).

Supported by a Cancer Center Support Grant (National Cancer Institute grant P30 CA016672).

Author Contributions: D.E.O. was the principal investigator (PI) for this study and was responsible for project conception, oversight, organization, data collection and auditing, statistical analysis, and manuscript writing. G.M.-Z. was involved in data collection, statistical analysis, and manuscript writing. S.M. was involved in statistical analysis and contributed to writing. J.S. and L.L. were the primary biostatisticians for the project, constructed models and analysis, and contributed to writing. F.A.A. (PI, Cleveland Clinic), M.J.S. (PI, Henry Ford Hospital), L.Y. (PI, Johns Hopkins Hospital), B.Y., D.F.-K., A.-E.S.S., T.G., L.G.D., H.B.G., R.F.C., G.A.E., and C.A.J. contributed to data collection and writing. L.Z.N., M.H.A., and S.B. were involved in data extraction.

Correspondence and requests for reprints should be addressed to David E. Ost, M.D., M.P.H., The University of Texas MD Anderson Cancer Center, Pulmonary Department, 1515 Holcombe Boulevard, Unit 1462, Houston, TX 77030. E-mail: [dost@mdanderson.org](mailto:dost@mdanderson.org).

This article has an online supplement, which is accessible from this issue's table of contents at [www.atsjournals.org](http://www.atsjournals.org).

Am J Respir Crit Care Med Vol 201, Iss 2, pp 212–223, Jan 15, 2020

Copyright © 2020 by the American Thoracic Society

Originally Published in Press as DOI: 10.1164/rccm.201904-0831OC on October 1, 2019

Internet address: [www.atsjournals.org](http://www.atsjournals.org)

## At a Glance Commentary

### Scientific Knowledge on the

**Subject:** A model previously created by O'Connell and colleagues, Help with Assessment of Adenopathy in Lung Cancer (HAL), uses histology, patient age, positron emission tomography–computed tomography N stage, and location of the tumor to estimate the probability of N2 or N3 malignant disease as determined by endobronchial ultrasound–guided transbronchial needle aspiration in patients with non–small cell lung cancer. However, for nonsurgical candidates in whom stereotactic ablative radiotherapy is a treatment option, it is important to distinguish between N0 and N1 disease because stereotactic ablative radiotherapy is not effective for N1 disease.

### What This Study Adds to the Field:

A model using histology, patient age, positron emission tomography–computed tomography N stage, and location of the tumor was able to accurately predict the probability of N0, N1, and N2 or N3 malignant nodal disease as determined by endobronchial ultrasound–guided transbronchial needle aspiration in patients with non–small cell lung cancer. The model was temporally and externally validated, demonstrating good discrimination. Calibration of the model was required for some external sites.

inoperable, or in patients who refuse surgery, stereotactic ablative radiotherapy (SABR) is recommended (2). For higher cancer stages, multimodality therapy with chemoradiation, chemotherapy, or targeted therapy is preferred (2, 3).

Previously, O'Connell and colleagues published a prediction model called Help with the Assessment of Adenopathy in Lung Cancer (HAL) (4). This model predicts the probability (pr) of having N2 or N3 (prN2|3) disease as determined by endobronchial ultrasound (EBUS)-guided transbronchial needle aspiration (EBUS-TBNA) in patients with NSCLC. In this model, younger age, central tumor location, adenocarcinoma, and higher N stage by

positron emission tomography (PET)–computed tomography (CT) (PET-CT) were all associated with increased prN2|3 disease versus N0 or N1 (N0|1) disease.

However, for patients being considered for SABR, it is important to distinguish between N0 and N1 disease because ablative radiation is directed only to the primary tumor without covering N1 nodes (5, 6). If N1 disease is present, SABR may not suffice (5, 6). Predicting the prN0, N1, and N2|3 nodal disease in patients with NSCLC requires a different model because HAL cannot distinguish between N0 and N1 disease. Accurate estimates of the prN0, N1, and N2|3 disease are central to the decision-making process and help to drive staging and treatment decision in patients with NSCLC in whom SABR is being considered (7, 8).

In this study, our objective was to create a prediction model for estimating the prN0, N1, and N2|3 lymph node involvement in patients with NSCLC. The secondary objective was to temporally and externally validate the model.

Some of the results of this study have been presented in abstract form during the American Thoracic Society 2019 International Conference (9).

## Methods

The development, temporal validation, and external validation cohorts shared the same inclusion and exclusion criteria. All consecutive untreated patients with NSCLC, clinical-radiographic stage T1 to T3, N0 to N3, and M0 who underwent EBUS-TBNA for staging were included. Patients with distant metastasis, mediastinal invasion by CT, suspected or confirmed synchronous primaries, recurrent lung cancer, and small cell cancer were excluded. Patients without PET imaging before treatment were excluded.

For the development cohort, we performed a retrospective analysis of consecutive patients with NSCLC who underwent staging EBUS-TBNA from September 2009 to January 2013. The study was approved by the Institutional Review Board Committee 5, Protocol PA16–0107, at The University of Texas MD Anderson Cancer Center. Data were collected prospectively as part of the American

College of Chest Physicians Quality Improvement Registry, Evaluation, and Education, as previously reported (4, 10–13). We used standardized definitions, quality control checks, and entered data into a Web-based interface (REDCap). This data set was the same one used to develop the HAL model (4).

PET-CT scan was used to define location of the tumor (central vs. peripheral) and N stage. Definitions of all variables were developed before data abstraction and provided to all sites. For CT scans, abnormal lymph nodes were defined as being  $\geq 1$  cm in their short axis. Lymph node N stage was determined by review of the radiology report and further review by an interventional pulmonologist or an interventional pulmonary fellow under supervision to assign an N stage to the patient. If both contrast and noncontrast CT were available, the contrast-enhanced images were used to determine CT N stage.

PET N stage was based on the radiologist's interpretation of mediastinal lymph node fluorodeoxyglucose F 18 avidity. In some cases standardized uptake value (SUV) measurements were recorded. In those cases that SUV measures on lymph nodes were available, an SUV value greater than or equal to 2.5 was considered positive. Based on the radiologist's reading and further review by an interventional pulmonologist or a supervised interventional pulmonary fellow, the PET N stage of the lesion was determined. Radiographic N stage by PET-CT was defined as the highest abnormal nodal station using *The Eighth Edition Lung Cancer Stage Classification* (14).

Tumors in the inner one-third of the hemithorax were defined as central (Figure E1 in the online supplement) (15, 16). All EBUS-TBNA procedures sampled N3 followed by N2 and then N1 nodes. All lymph nodes measuring 0.5 cm or larger by EBUS were sampled, independent of PET-CT status.

## Statistical Analysis

**Prediction model development.** The primary outcome was the highest N stage (N0 vs. N1 vs. N2|3) lymph node with malignancy as determined by EBUS-TBNA. N0, N1, and N2|3 disease groups were compared using the Fisher exact test for categorical variables and ANOVA for continuous normally distributed variables. Because PET-CT

**Table 1.** Descriptive Statistics by N Stage in the Development Cohort (N = 633)

	<i>n</i> Missing	N0 ( <i>n</i> = 412)	N1 ( <i>n</i> = 61)	N2 3* ( <i>n</i> = 160)	<i>P</i> Value <sup>†</sup>
Age, yr, mean ± SD	0	68.99 ± 9.3	66.57 ± 10.01	65.23 ± 10.49	0.001 <sup>‡</sup>
Sex, <i>n</i> (%)	0				
F		194 (63.6)	28 (9.2)	83 (51.9%)	0.549
M		218 (66.5)	33 (10.1)	77 (23.5)	
Race, <i>n</i> (%)	4				0.922
Asian		13 (76.5)	2 (11.8)	2 (11.7)	
Black		31 (64.6)	4 (8.3)	13 (27.1)	
Hispanic		20 (66.7)	3 (10.0)	7 (23.3)	
White		347 (65.0)	50 (9.4)	137 (25.7)	
ASA score, <i>n</i> (%)	0				0.814
1		3 (50.0)	0 (0.0)	3 (50.0)	
2		31 (63.3)	4 (8.2)	14 (28.6)	
3		372 (65.1)	57 (10.0)	142 (24.9)	
4		6 (85.7)	0 (0.0)	1 (14.3)	
Smoking status, <i>n</i> (%)	0				0.748
Current smoker		89 (66.4)	11 (8.2)	34 (25.4)	
Never smoker		39 (66.1)	8 (13.6)	12 (20.3)	
Prior smoker		284 (64.5)	42 (9.5)	114 (25.9)	
ECOG, <i>n</i> (%)	0				0.181
0		112 (64.3)	23 (13.2)	39 (22.4)	
1		210 (62.9)	28 (8.4)	96 (27.7)	
2		76 (70.4)	9 (60.8)	23 (21.3)	
3		14 (82.4)	1 (5.9)	2 (11.8)	
Size of the tumor, <i>n</i> (%)	0				0.298
≤3 cm		186 (68.4)	23 (8.5)	63 (23.2)	
>3 cm but ≤5 cm		132 (62.9)	18 (8.6)	60 (28.6)	
>5 cm		94 (62.3)	20 (13.2)	37 (24.2)	
Lobar location of the tumor, <i>n</i> (%)	0				0.386
Left upper lobe or lingula		123 (68.0)	18 (9.9)	40 (22.1)	
Left lower lobe		66 (70.2)	8 (8.5)	20 (23.8)	
Right upper lobe		140 (65.7)	18 (8.5)	55 (25.8)	
Right lower or middle lobe		83 (57.2)	17 (11.7)	45 (31.0)	
Location, <i>n</i> (%)	0				0.039
Outer two-thirds of lung		323 (78.4)	42 (68.9)	111 (69.4)	
Central one-third of lung		89 (21.6)	19 (31.1)	49 (30.6)	
Histology, <i>n</i> (%)	0				0.054
Adenocarcinoma		203 (61.3)	31 (9.4)	97 (29.3)	
Squamous cell carcinoma		158 (72.8)	19 (8.8)	40 (18.4)	
Non-small cell carcinoma		32 (55.2)	8 (13.2)	18 (31.0)	
Other primary lung cancer		19 (70.4)	3 (11.1)	5 (18.5)	
CT characteristics, <i>n</i> (%)	0				0.016
Cavitary		15 (68.2)	3 (13.6)	4 (18.2)	
Ground glass/semisolid/infiltrate		32 (88.9)	0 (0.0)	4 (11.1)	
Solid		365 (63.5)	58 (10.1)	152 (26.4)	
Satellite lesion in same lobe, <i>n</i> (%)	0				0.160
No		392 (95.1)	60 (98.4)	157 (98.1)	
Yes		20 (4.9)	1 (1.6)	3 (1.9)	
N stage by PET-CT, <i>n</i> (%)	0				<0.001
CT = N0 1; PET = N0		171 (95.0)	2 (1.1)	7 (3.8)	
CT = N2 3; PET = N0		79 (86.8)	4 (4.4)	8 (8.8)	
CT = N0 1; PET = N1		38 (48.1)	28 (35.4)	13 (16.5)	
CT = N2 3; PET = N1		19 (48.7)	16 (41.0)	4 (10.3)	
CT = N0 1; PET = N2 3		44 (68.8)	2 (3.1)	18 (28.1)	
CT = N2 3; PET = N2 3		61 (33.9)	9 (5.0)	110 (61.1)	

*Definition of abbreviations:* ASA = American Society of Anesthesiologists; ECOG = Eastern Cooperative Oncology Group; CT = computed tomography; N0|1 = N0 or N1; N2|3 = N2 or N3; PET = positron emission tomography.

N stage as assessed by endobronchial ultrasound-guided transbronchial needle aspiration.

\*N2 and N3 are combined (N2|3).

<sup>†</sup>*P* values are for chi-square test except where otherwise noted.

<sup>‡</sup>ANOVA.

images do not use contrast, and some patients only had noncontrast CT available, we used N0|1 disease as a single variable for CT but they were kept separate for PET,

as previously reported (*see* online supplement) (4).

We used univariate ordinal logistic regression to identify variables associated

with the outcome variable, highest N stage by EBUS, classified in the following order: N0 < N1 < N2|3. With three ordinal outcomes, two sets of probabilities are

calculated. The first is the prN stage being greater than or equal to 1 (prN1|2|3) versus the prN0 disease. The second is the prN stage being N2|3 versus N0|1. We specified *a priori* that variables with an overall *P* value less than 0.2 on univariate analysis would be candidate variables for the multivariable ordinal logistic regression model. We checked the proportional odds assumption using the Score test. Different slope parameters were allowed for variables that violated this assumption (*see* online supplement). We specified *a priori* that we would use stepwise backward selection with an overall *P* value less than 0.05 for variables to remain in the model.

### Temporal and External Validation

For temporal validation, data from a completely different cohort of MD Anderson Cancer Center patients who underwent EBUS-TBNA from September 2016 to January 2019 were used. These data were prospectively collected and constituted the temporal validation cohort (17, 18).

For external validation, data from three centers (Johns Hopkins, Henry Ford Hospital, and Cleveland Clinic Foundation) were used (17, 18). Consecutive patients

were entered using identical definitions, forms, and quality control checks as in the development cohort. These patients constitute the external validation cohort and correspond to the external validation cohort in the HAL model (4).

### Model Performance Assessment

We assessed model performance in the development, temporal validation, and external validation cohorts. We used the receiver operating characteristic (ROC) area under the curve (AUC) to assess discrimination. We used the Hosmer-Lemeshow goodness-of-fit test and Brier score, and observed versus predicted graphs, to assess calibration (*see* online supplement).

We created a calibrated model for the combined data from all three outside institutions and a separate calibrated model for each institution using the general calibration method presented by Steyerberg and colleagues (*see* online supplement) as previously reported in the HAL model (*see* online supplement for additional details) (4, 19).

For the temporal validation cohort, we hypothesized that the model would not require further calibration because the location was the same as the development

cohort and we wanted to test model stability over time. Therefore, we prespecified that we would not calibrate the temporal validation cohort whatsoever, and measured discrimination and calibration using the baseline model.

All statistical analyses used SAS version 9.4 (SAS Institute) or STATA 15.1 (StataCorp LLC).

## Results

The development cohort consisted of 633 patients. Descriptive statistics for the cohort stratified by final EBUS N stage are in Table 1.

### Model Development

Univariate ordinal logistic regression results are in Table E1 and multivariate results are in Table 2. The only candidate variable that violated the proportional odds assumption was N stage by PET-CT (*P* < 0.001). For this variable, different slope parameters were allowed. Younger age, adenocarcinoma histology, central location, and higher nodal stage by PET-CT were associated with an increased prN1|2|3 (vs. prN0) disease and higher prN2|3 (vs. prN0|1) disease (*see* Table 2). ROC AUC

**Table 2.** Multivariate Ordinal Logistic Regression Model for Prediction of N0 versus N1 versus N2|N3 Disease

	N1 2 3 (vs. N0) Disease				N2 3 (vs. N0 1) Disease			
	Coefficient	Odds Ratio	95% CI	<i>P</i> Value	Coefficient	Odds Ratio	95% CI	<i>P</i> Value
Age, yr*	-0.029	0.97	0.95-0.99	0.003	-0.029	0.97	0.95-0.99	0.003
Tumor location*								
Outer two-thirds of the lung	0	1.00	—	—	0	1.00	—	—
Central one-third of the lung	0.486	1.62	1.03-2.55	0.034	0.486	1.62	1.03-2.55	0.034
Tumor histology*								
Adenocarcinoma	0	1.00	—	—	0	1.00	—	—
Squamous-cell carcinoma	-0.821	0.44	0.28-0.68	<0.001	-0.821	0.44	0.28-0.68	<0.001
Non-small cell lung carcinoma	0.063	1.06	0.55-2.03	0.847	0.063	1.06	0.55-2.03	0.847
Other primary	-0.409	0.66	0.25-1.75	0.409	-0.409	0.66	0.25-1.75	0.409
N stage by PET-CT†								
CT = N0 1; PET = N0	0	1.00	—	—	0	1.00	—	—
CT = N2 3; PET = 0	1.173	3.23	1.29-8.08	0.012	0.979	0.97	0.92-7.67	0.069
CT = N0 1; PET = N1	3.083	21.82	9.62-49.48	<0.001	1.593	1.59	1.85-13.03	0.001
CT = N2 3; PET = N1	2.990	19.89	7.76-50.93	<0.001	0.932	0.93	0.69-9.24	0.157
CT = N0 1; PET = N2 3	2.259	9.57	4.00-22.88	<0.001	2.359	2.35	4.10-27.32	<0.001
CT = N2 3; PET = N2 3	3.711	40.90	19.25-86.90	<0.001	3.748	3.74	18.58-97.06	<0.001
Constant‡	-0.890	—	—	0.233	-1.1576	—	—	0.131

*Definition of abbreviations:* CI = confidence interval; CT = computed tomography; N0|1 = N0 or N1; N1|2|3 = N stage greater than or equal to 1; N2|3 = N2 or N3; PET = positron emission tomography.

\*Variables did not violate the proportional odds assumption in the univariate analysis. Therefore, the coefficients for N1|2|3 (vs. N0) disease are the same as the coefficients of N2|3 (vs. N0|1) disease.

†Variable violated the proportional odds assumption in the univariate analysis. Therefore, two slope parameters were obtained, one for N1|2|3 (vs. N0) disease and one for N2|3 (vs. N0|1) disease. Both coefficients are shown.

‡Two constants were calculated: one for the formula used to predict N1|2|3 (vs. N0) disease and one for the formula used to predict N2|3 (vs. N0|1) disease.

was 0.84 (95% confidence interval [CI], 0.81–0.87) for predicting N1|2|3 (vs. N0) disease and 0.85 (95% CI, 0.82–0.89) for predicting N2|3 (vs. N0|1) disease (Figures 1A and 1B). Model fit was acceptable (Hosmer-Lemeshow test,  $P = 0.960$ ; Brier score, 0.36; observed vs. predicted graphs) (Figures 2A and 2B).

### Temporal Validation

The temporal validation cohort included 473 patients (see Table E2). ROC AUC was 0.86 (95% CI, 0.85–0.90) for predicting N1|2|3 (vs. N0) disease and 0.88 (95% CI, 0.84–0.92) for predicting N2|3 (vs. N0|1) disease (see Figures 1C and 1D).

Model fit was acceptable (Hosmer-Lemeshow test,  $P = 0.172$ ; Brier score, 0.30; observed vs. predicted graphs) (see Figures 2C and 2D). There was no need to calibrate the model, suggesting that the model is

stable and accurate over time in this location.

### External Validation

The external validation cohort included 722 patients (see Table E3). Discrimination was good for the combined external validation cohort and for each outside institution when assessed separately. For the combined external validation cohort, AUC was 0.86 (95% CI, 0.84–0.89) for predicting N1|2|3 (vs. N0) disease and 0.88 (95% CI, 0.85–0.90) for predicting N2|3 (vs. N0|1) disease (see Figure E2). The AUCs for each outside institution ranged from 0.81 to 0.91 for predicting N1|2|3 (vs. N0) disease and from 0.82 to 0.92 for predicting N2|3 (vs. N0|1) disease (Table 3).

When assessing calibration of the combined external validation cohort, model fit was not acceptable (Hosmer-Lemeshow,

$P < 0.001$ ; Brier score, 0.38; observed vs. predicted graphs) (Figures 3A and 3B). When we assessed model calibration in each institution separately, model fit was acceptable in one of the outside institutions (Hosmer-Lemeshow,  $P = 0.286$ ; Brier score, 0.34; observed vs. predicted graphs) (Figures 4A and 4B; see Table 3) but was off in two of the external validation sites (Hosmer-Lemeshow,  $P < 0.001$ ; Brier score range 0.36–0.46; observed vs. predicted graphs) (see Figures 4C and 4F; see Table 3).

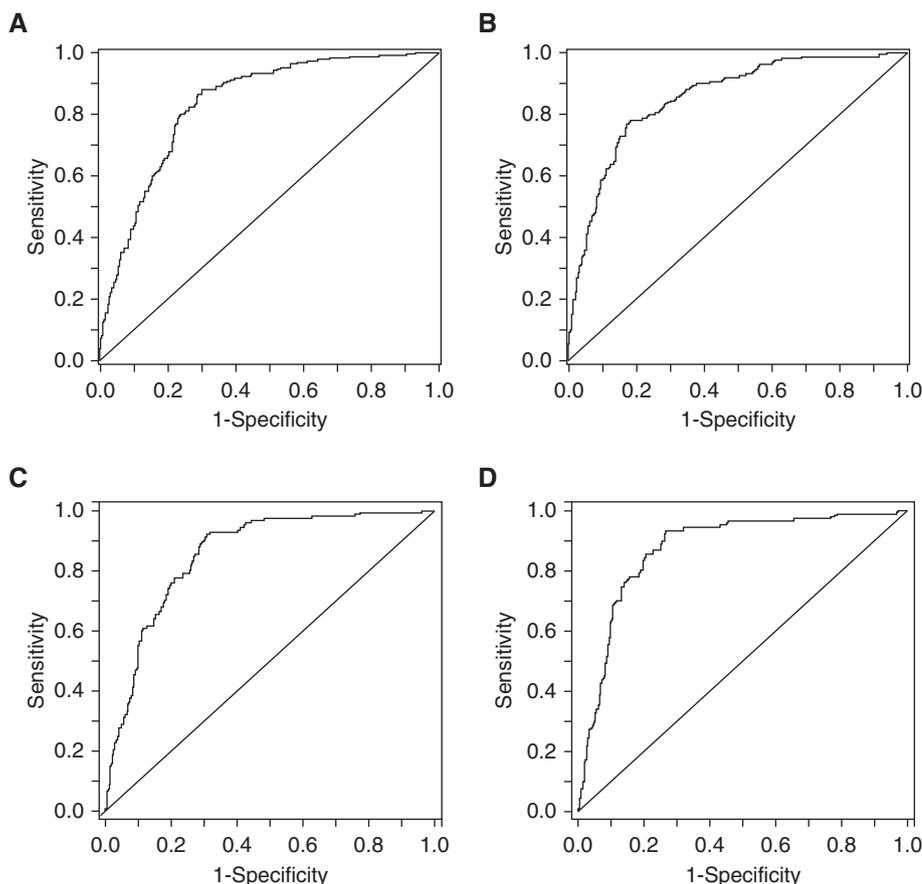
### Calibration of the External Validation Cohorts

When we calibrated the model to predict outcomes for the combined external validation cohort, two sets of slope and intercept were calculated, one for predicting N1|2|3 (vs. N0) disease and another for N2|3 (vs. N0|1) disease. Both calibration intercepts were the same (0.75), but the slopes were off by 0.01 (1.13 vs. 1.14, respectively). Both calibrations were evaluated (intercept = 0.75; slopes = 1.13 and 1.14), and the one with lower Brier score was selected (intercept = 0.75; slope = 1.14). After calibration, the Hosmer-Lemeshow test was nonsignificant ( $P = 0.094$ ), and both the Brier score (0.34) and observed versus predicted graphs (see Figures 3C and 3D) showed improved model fit.

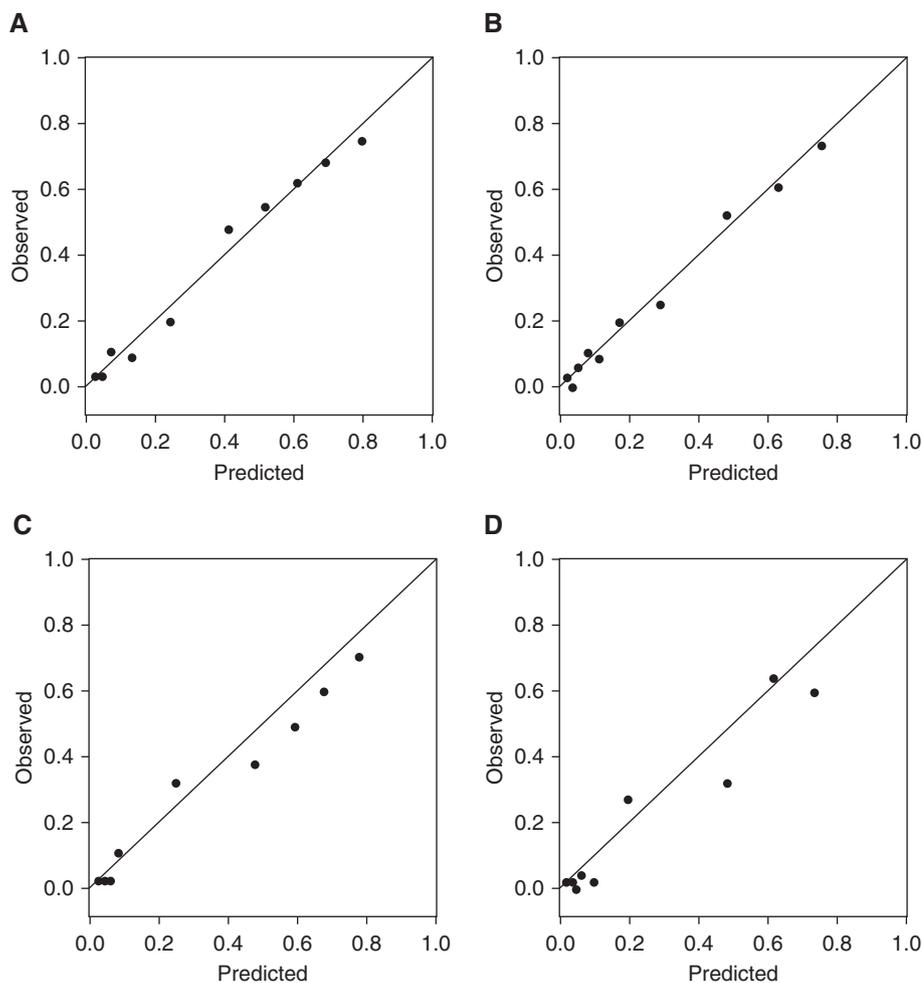
For model calibration of each outside institution, the slope for all three centers was set to unity ( $b = 1$ ). The pair of intercepts with minimum Brier score and maximum Hosmer-Lemeshow  $P$  value was selected (see Table E4). The institution-specific calibrated models performed well. Hosmer-Lemeshow tests became nonsignificant ( $P = 0.196$ – $0.404$ ) (see Table E4), Brier scores improved (range, 0.29–0.39) (see Table E4), and observed versus predicted graphs showed improved model fit (Figure 5).

### Discussion

In this study, we report on the Help with Oncologic Mediastinal Evaluation for Radiation (HOMER) model, which estimates the prN0 versus N1 versus N2|3 metastatic nodal disease in patients with NSCLC. We demonstrated that HOMER is accurate in outside institutions after calibration using the general calibration method. We also demonstrated that the



**Figure 1.** Receiver operating characteristic curves of the prediction model in the institution of model development. The figure plots the area under the curve (AUC) for (A) N stage greater than or equal to 1 (vs. N0) disease (AUC = 0.84) and (B) N2 or N3 (vs. N0 or N1) disease (AUC = 0.85) in the development cohort, and for (C) N stage greater than or equal to 1 (vs. N0) disease (AUC = 0.86) and (D) N2 or N3 (vs. N0 or N1) disease (AUC = 0.88) in the temporal validation cohort.



**Figure 2.** Observed versus predicted frequencies of the prediction model in the institution of model development. The figure plots the probability of (A) N stage greater than or equal to 1 (vs. N0) disease and (B) N2 or N3 (vs. N0 or N1) disease by decile of expected risk in the group of the development cohort, and the probability of (C) N stage greater than or equal to 1 (vs. N0) disease and (D) N2 or N3 (vs. N0 or N1) disease by decile of expected risk in the group of the temporal validation cohort. The observed probability for each decile is on the vertical axis, and the predicted probability is on the horizontal axis. A perfect model, in which observed equals predicted, is shown by the line.

model maintained good discrimination and calibration over an extended period of time when applied to a single institution using two different data sets.

The American College of Chest Physicians and National Comprehensive Cancer Network lung cancer guidelines suggest using prediction models to estimate

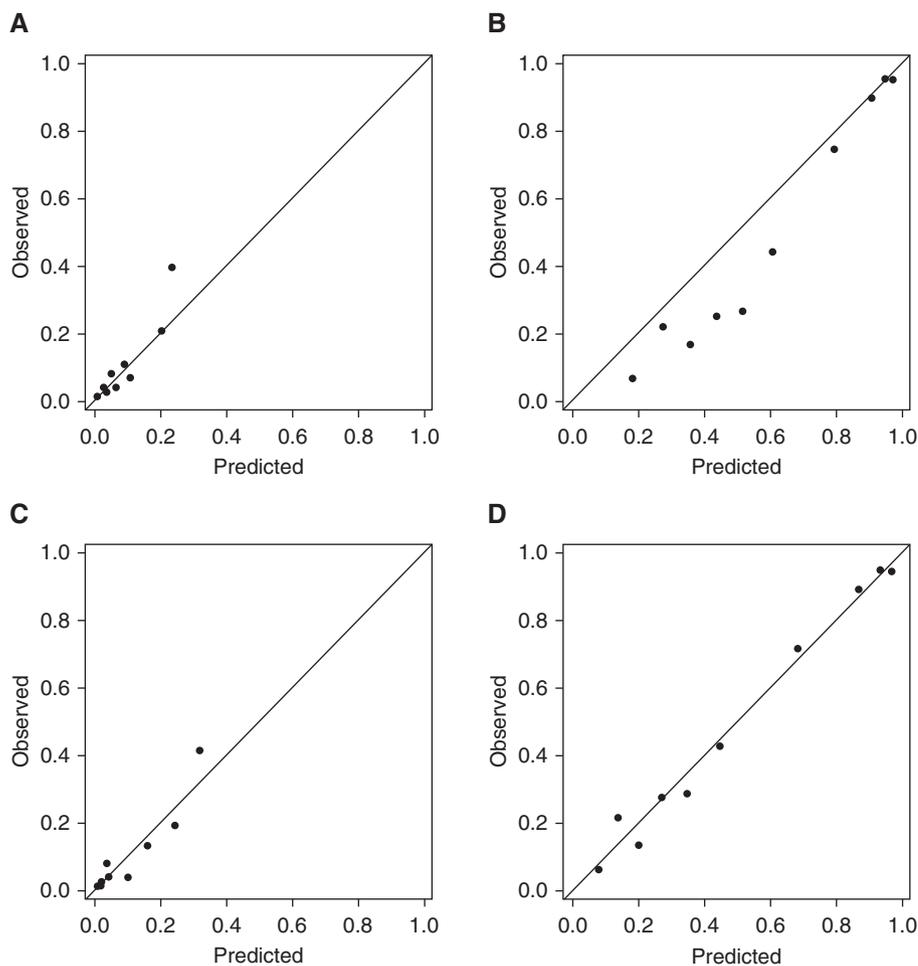
the probability of malignancy in solitary pulmonary nodules to help inform decision-making in patients with solitary pulmonary nodules (2, 20, 21). Similarly, investigators have developed binary prediction models to estimate prN2|3 (vs. prN0|1) to help inform decision-making in patients with NSCLC with regard to staging procedures when surgical treatment is the main option (4, 22–26). However, those studies did not distinguish between N0 and N1 disease, which is critical when SABR is a treatment option. One study did use separate binary logistic regression models to identify risk factors for N1 versus N0 disease and for N2 versus N0 disease in patients who had surgery (27). However, that study did not include nonsurgical patients, limiting the generalizability of the findings. More importantly, because the investigators used multiple binary logistic regression models rather than a single ordinal logistic regression model, their model is not valid for clinical prediction. That is because the form of their models is 1) given that a patient has N0|1 disease, then the odds are X and 2) given that a patient has N0 or N2 disease, then the odds are Y. In real life, physicians cannot know *a priori* that N1, N2, or N3 disease is definitely absent. Therefore, multiple binary models of this form cannot work for clinical prediction.

This study adds to the existing body of knowledge by using ordinal logistic regression to develop a more generalizable prediction rule to inform decision-making in patients who are candidates for SABR. To our knowledge, HOMER is the first externally and temporally validated model to predict prN0, prN1, and prN2|3 disease developed from a broad population of patients that included both surgical and nonsurgical candidates. It incorporates PET imaging, which is part of the current standard of care but does not rely on molecular markers. Because it includes both

**Table 3.** Model Performance at Outside Institutions: Predictions before Calibration

Institution	Brier Score	Hosmer-Lemeshow ( <i>P Value</i> )	AUC (95% CI)	
			N1 2 3 (vs. N0) Disease	N2 3 (vs. N0 1) Disease
CCF (N=310)	0.34	0.286	0.84 (0.79–0.88)	0.87 (0.83–0.91)
JH (N=186)	0.46	<0.001	0.82 (0.74–0.89)	0.82 (0.76–0.89)
HFH (N=226)	0.36	<0.001	0.91 (0.87–0.95)	0.92 (0.88–0.95)

*Definition of abbreviations:* AUC = area under the receiver operating characteristics curve; CI = confidence interval; CCF = Cleveland Clinic Foundation; JH = Johns Hopkins; HFH = Henry Ford Hospital; N = nodal stage; N0|1 = N0 or N1; N1|2 = N1 or N2; N1|2|3 = N stage greater than or equal to 1; N2|3 = N2 or N3.



**Figure 3.** Observed versus predicted frequencies for combined external validation cohort. The figure plots the probability of (A) N stage greater than or equal to 1 (vs. N0) disease and (B) N2 or N3 (vs. N0 or N1) disease by decile of expected risk in that group before calibration, and the probability of (C) N stage greater than or equal to 1 (vs. N0) disease and (D) N2 or N3 (vs. N0 or N1) disease by decile of expected risk in that group after calibration. The observed probability for each decile is on the vertical axis, and the predicted probability is on the horizontal axis. A perfect model, in which observed equals predicted, is shown by the line.

surgical and nonsurgical candidates, the model is applicable to patients in whom SABR is the only option and to patients in whom SABR is being considered as an alternative to surgery (e.g., borderline surgical candidates with T1b, N0, and M0 disease by PET-CT) (2, 8).

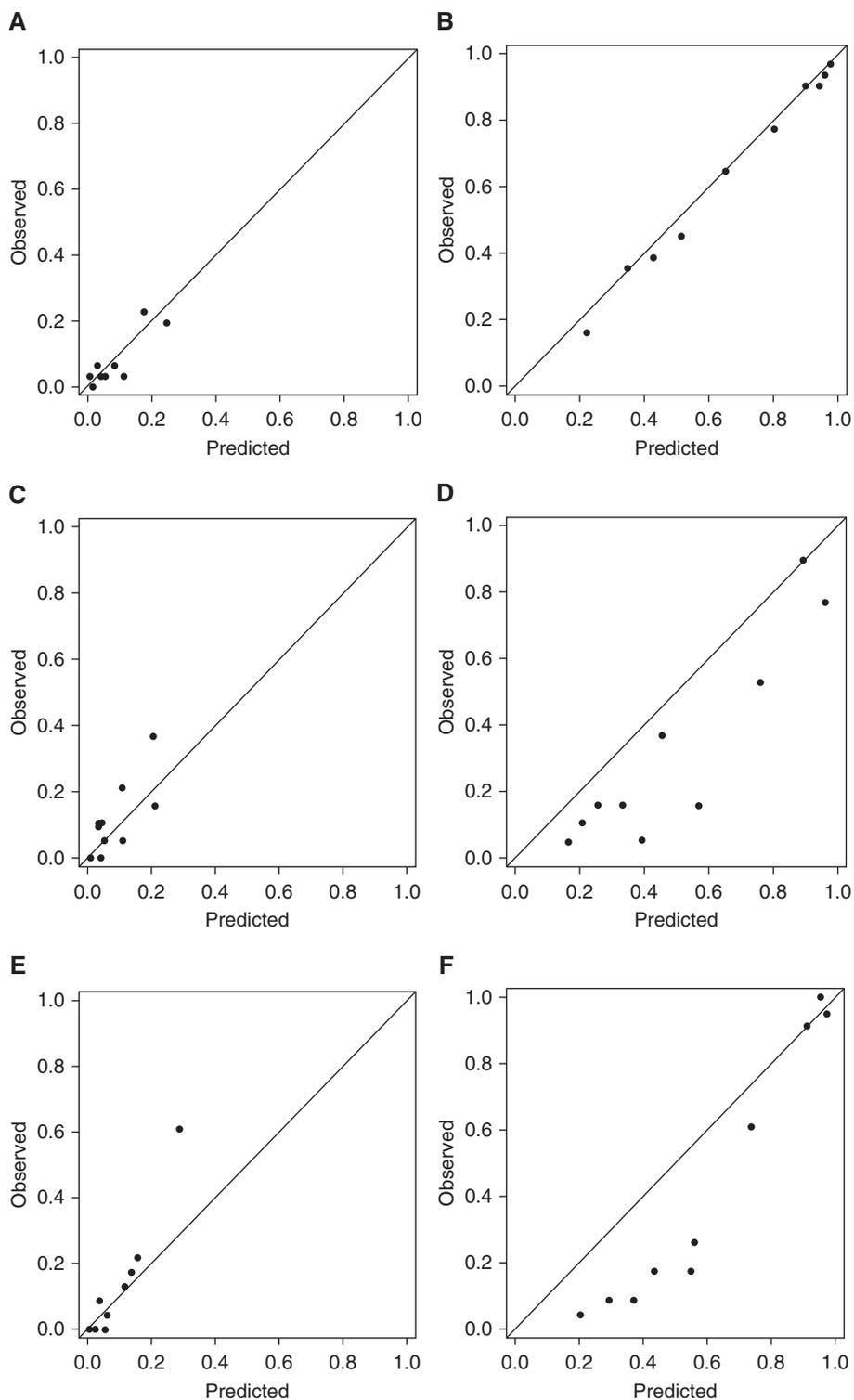
Although there is currently an ongoing prospective study of endosonographic intrathoracic nodal staging of patients being considered for therapy with SABR, there are currently no definitive recommendations on whether to perform EBUS for mediastinal staging before SABR (28, 29). The role of EBUS before SABR depends in large part on the context. EBUS could be useful to inform decisions when patients are candidates for both surgery and SABR (e.g.,

PET-CT N0) (30). A finding of N1 disease would lead to a clear recommendation in such cases. Whether EBUS should be done in this context is a function of the probability of EBUS being positive and the complication rate of EBUS (12). Given the low rate of complications from EBUS, even a relatively low prN1 disease in these patients might warrant EBUS. Conversely, in patients who are not surgical candidates, the need for EBUS is different. Finding N1 disease in such patients would probably lead to definitive radiation therapy, whereas finding N2 disease would lead to multimodal treatment with radiation and chemotherapy instead of treatment with SABR. Whether EBUS is warranted in this context would be a function of the

probability of EBUS being positive, the complication rate of EBUS, and the marginal benefit and marginal harm of treating with chemotherapy and radiation for N2 disease; and definitive radiation for N1 disease (as compared with treating with SABR based solely on imaging) (31–49). However, in both cases, accurately predicting the probability of EBUS being positive is vital. HOMER can aid in this decision-making process.

Consideration of specific cases may help illustrate these concepts. Consider a 60-year-old patient with adenocarcinoma in the outer two-thirds of the lung, PET-CT N0, who is a surgical candidate but prefers not to have surgery if possible, and SABR is being considered. HOMER predicts prN0 is 93% by EBUS (see Table E5). When the physician weighs the risk of complications at approximately 1.15% versus a 7% chance of having prN1|2|3 and changing treatment from SABR to surgery, EBUS seems warranted (12). However, in other circumstances, HOMER might lead to a decision not to do EBUS. Consider an 80-year-old patient with squamous cell carcinoma in the outer two-thirds of the lung, PET-CT N0, who is not a candidate for surgery. HOMER predicts prN0 is 98% by EBUS (see Table E5). Given the risks of EBUS, the marginal benefit and marginal harm of treating with chemotherapy and radiation instead of SABR for occult N2 disease, and only a 2% chance of EBUS being positive, proceeding directly to SABR is reasonable (31–49). The absolute difference in prN0 in these two patients is only 5%, which at first glance seems small. However, given the low risk of EBUS and consideration of the benefits and harms of treatment, the value of information in this context is high (7, 8, 50).

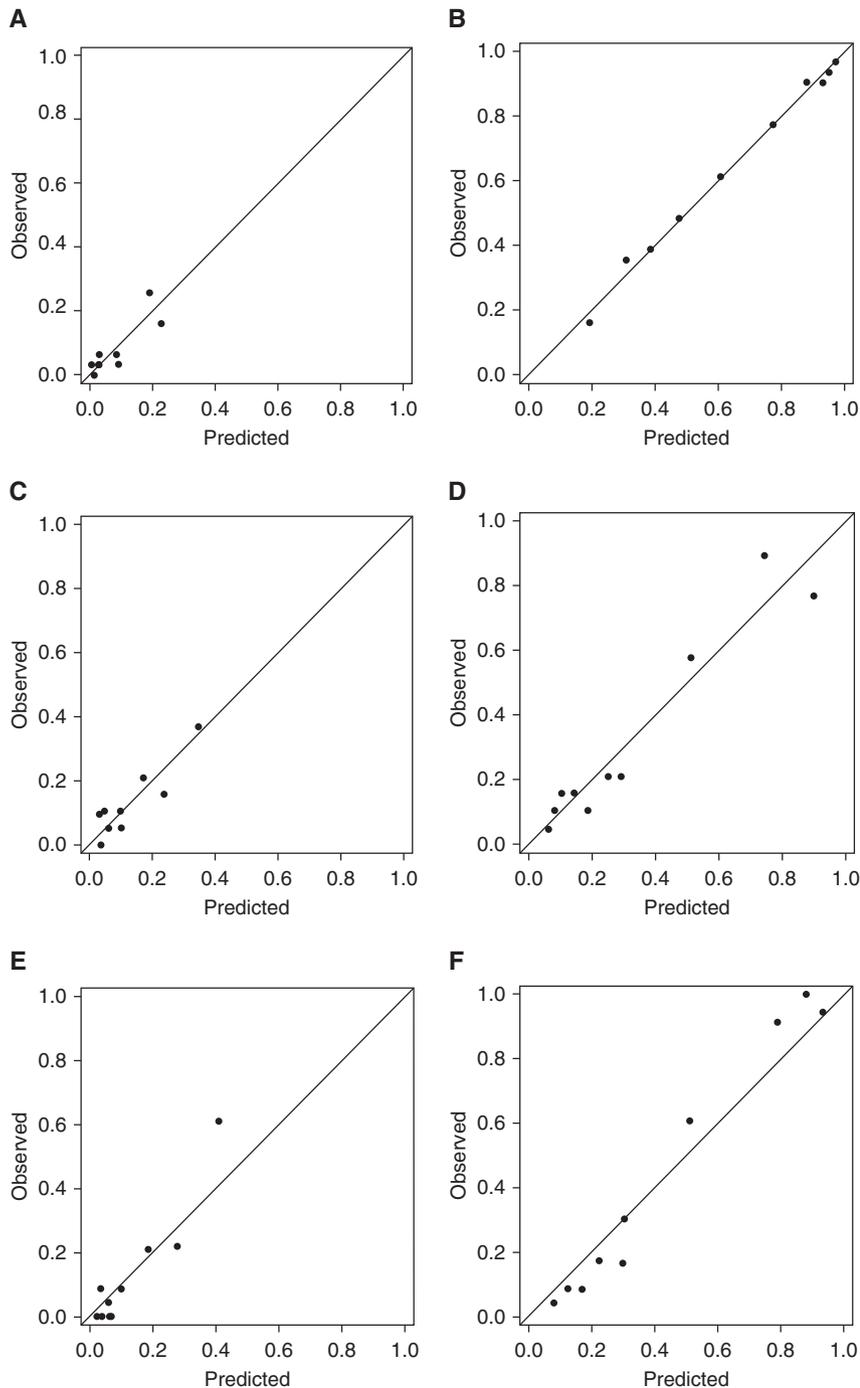
Another practical application of HOMER is helping to inform decisions in patients with clinical-radiographic N1 disease. Previous studies have pointed out that EBUS in this population may downstage them, making them potentially suitable for SABR therapy (6). If EBUS is performed and the patient is N0 by EBUS, the decision on whether SABR is a reasonable choice depends on knowledge of the posttest probability of nodal disease, as well as the benefits and harms of SABR versus other treatment alternatives. The probability of nodal disease after a negative EBUS is of course a function of the



**Figure 4.** Observed versus predicted frequencies for each institution of the external validation cohort before calibration. The figure plots the probability of (A) N stage greater than or equal to 1 (N1|2|3) (vs. N0) disease and (B) N2 or N3 (N2|3) (vs. N0 or N1 [N0|1]) disease at the Cleveland Clinic Foundation, the probability of (C) N1|2|3 (vs. N0) disease and (D) N2|3 (vs. N0|1) disease at Johns Hopkins, and the probability of (E) N1|2|3 (vs. N0) disease and (F) N2|3 (vs. N0|1) disease at the Henry Ford Hospital. The observed probability for each decile is on the vertical axis, the predicted probability on the horizontal axis. A perfect model, in which observed equals predicted, is shown by the line.

sensitivity of EBUS. The sensitivity of EBUS varies with PET-CT stage (6, 51–64). If EBUS sensitivity is 0.8 in the setting of N1 disease by PET-CT, then we can use HOMER to estimate the posttest probability that a patient with a negative EBUS actually has N0, N1, or N2|3 disease. In a 60-year-old patient with adenocarcinoma in the outer two-thirds of the lung, PET-CT N1 disease, assuming EBUS sensitivity is 0.8, the probability of true N0 disease given a negative EBUS is only 61% (*see* online supplement for calculations). If EBUS sensitivity is 0.9, the probability of true N0 disease is 83%. Conversely, if the patient is 80 years old with a peripheral squamous cell in the outer two-thirds of the lung, PET-CT N1, with a negative EBUS, then the corresponding probabilities for true N0 disease are 91% and 96%. In scenarios in which EBUS is negative but HOMER predicts a 17% to 39% (or a 4–9% in the second scenario) probability of having nodal metastasis, other factors related to the benefit and harm of SABR alone versus alternative strategies become relevant. These other factors include the performance status of the patient, tumor marker status, and the multimodality options being considered. If definitive radiotherapy covering the hilar nodes is used, toxicity will be higher but there will be the possibility of cure. Conversely, if SABR is used and targeted therapy is used to treat subsequent relapses, toxicity will be lower but the treatment will be palliative in nature. By providing the predicted probability, HOMER can help inform this decision, adding nuance to this complex decision process.

By using an ordinal model, HOMER also provides additional insights into the relationship of the predictor variables to N stage that are lacking in binary models and that might fail to capture the entire complexity of a clinical decision problem owing to simplification and loss of information. HOMER demonstrates that, for older patients with N2|3 disease by PET, the most likely stage by EBUS is not necessarily N2|3 (and not N1), but rather it can be either N2|3 or N0 disease, depending on tumor location (*see* Figure E3C). For younger patients with N2|3 disease by PET, the most likely stage by EBUS is N2|3, with N0 disease being a close second. Quantifying these probabilities using HOMER (*see* Figure E3C) shows us that N1 disease by EBUS is actually a rare finding in PET-CT N2|3



**Figure 5.** Observed versus predicted frequencies for each institution of the external validation cohort after calibration. The figure plots the probability of (A) N stage greater than or equal to 1 (N1|2|3) (vs. N0) disease and (B) N2 or N3 (N2|3) (vs. N0 or N1 [N0|1]) disease at the Cleveland Clinic Foundation, the probability of (C) N1|2|3 (vs. N0) disease and (D) N2|3 (vs. N0|1) disease at Johns Hopkins, and the probability of (E) N1|2|3 (vs. N0) disease and (F) N2|3 (vs. N0|1) disease at the Henry Ford Hospital. The observed probability for each decile is on the vertical axis, the predicted probability on the horizontal axis. A perfect model, in which observed equals predicted, is shown by the line.

patients, with prN1 being much lower than either prN0 or prN2|3. No binary model can capture these subtleties (*see online supplement for further discussion*).

As in the HAL model, our study failed to demonstrate a relationship between tumor size and higher N stage by EBUS-TBNA after adjusting for age, tumor

location, tumor histology, and PET-CT N stage (4). This contrasts with other studies that have reported an association between larger tumors and probability of nodal metastatic disease (22–25). However, none of the prior studies adjusted for PET-CT N stage in their analysis. PET-CT N stage in this dataset is associated with tumor size ( $P=0.036$ ) (*see online supplement*). Therefore, PET-CT N stage potentially confounds the relationship between tumor size and nodal metastatic disease. This could explain the discordance in findings between studies. To make sure tumor size did not improve model performance, we forced tumor size back into the model, which did not improve discrimination but worsened calibration (*see online supplement*).

Effective prediction models should be validated on external cohorts and should demonstrate both good discrimination and calibration (*see online supplement*) (17, 18). HOMER demonstrates good discrimination in both the combined external validation cohort and for each institution when assessed separately. However, observed versus predicted graphs show that, although prediction of prN1|2|3 (vs. prN0) disease is decent, HOMER overestimates the prN2|3 (vs. prN0|1) disease for two out of the three outside institutions (*see Figures 3A, 3B, and 4*). Application of the general calibration method corrects this (*see Figures 3C, 3D, and 5*).

In regard to the observed versus predicted graphs for the temporal validation cohort, HOMER slightly overestimates prN1|2|3 (vs. prN0) disease when predicted probability is greater than 50% (*see Figures 2C and 2D*). Nonetheless, the Brier score and Hosmer-Lemeshow  $P$  values both show that the model has acceptable fit and does not require further calibration, suggesting that the model has temporal stability. Model stability over time is an important consideration. If the model is temporally stable, then calibration intercepts can be determined for any outside institution and from that point forward the model can make accurate predictions for patients at that institution.

To our knowledge, this is the first study of EBUS-TBNA to externally and temporally validate a prediction model for N0, N1, and N2|3 disease. The HAL model was the first study of EBUS-TBNA to externally validate a prediction model for N2|3 disease (4). In this study, we

externally validate HOMER, but we also temporally validate it, showing that it is stable through time in the institution of model development.

A limitation to the prediction power of our model is the small number of patients who fall in the least common of the three possible outcomes (N1 disease). The proportion of patients identified as having N1 disease by EBUS-TBNA in our data ranges from 7.2% to 10.11% for the development, temporal validation, and external validation cohorts. Our data are consistent with the findings of other investigators, in which the proportion of patients with N1 disease ranges from 5.3% to 16% (53, 54, 65, 66). The least common outcome determines the number of covariates a model can support, so having relatively little N1 disease limits the number of covariates that can be included for model development. Larger cohorts would be required if more covariates were to be introduced. It is possible that a larger sample size would allow identification of additional covariates, which might significantly improve model performance. Furthermore, the validity of the predictions made by HOMER are limited to patients who are between 60 to 80 years of age, the range of data in which 68% percent of patients from the development cohort are found (*see online supplement for details*). The model might not perform as well at the extremes of age.

Another limitation is that our predictions are for the observed cytology as identified by EBUS-TBNA, which itself has a

varying sensitivity that ranges from 54.5% to 98%, with a recent meta-analysis suggesting a pooled sensitivity of 90%; therefore, the model might underestimate the presence of true lymph node metastases (1, 52, 58, 66–69). Because HOMER does not predict the probabilities of N0, N1, or N2/3 disease as determined by thoracotomy, adjustment for EBUS sensitivity may be required, depending on how the model is used. A study of surgical patients undergoing thoracotomy would facilitate prediction of the true pretest probability of nodal disease and would help determine the sensitivity and specificity of EBUS-TBNA, but such a design would also have problems with generalizability and possibly selection bias because nonsurgical patients going for SABR and those with significant comorbidities could not be included. In addition, predicting pretest probability would not be sufficient by itself to determine whether EBUS-TBNA should be done in a given patient. What we need to know is the chance that EBUS-TBNA will be positive in that same patient (i.e., diagnostic yield). HOMER predicts the diagnostic yield of EBUS-TBNA for N0, N1, or N2/3 disease, which is fundamentally different than predicting the pretest probability of disease as determined by the gold standard (*see online supplement*). An additional limitation is that our prediction rule is only valid in centers that perform EBUS-TBNA in a similar systematic manner (4). All three centers of the external validation cohort are high-volume centers. Higher procedural volume is associated with higher diagnostic yield, so results may differ for

centers with lower procedural volume (13). A final limitation is that the Brier score used to assess calibration does not take into account the ordinal nature of the outcomes (*see online supplement*).

## Conclusions

The HOMER model predicts the probability of finding N0 versus N1 versus N2/3 disease on EBUS, and had good performance as assessed by tests of discrimination and calibration in the development cohort. The predictor variables identified were consistent with the previously reported HAL model (4). In regard to external validation, the model has good discrimination but requires calibration. After calibration, the model demonstrates sufficient precision to be useful clinically. Performance in the temporal validation cohort suggests that the model is stable through time in the institution where it was developed. The HOMER model is potentially useful for predicting N stage, and informing decisions regarding staging and treatment for patients with NSCLC in which SABR is an option. Future studies will need to assess whether calibrated models in outside institutions are temporally stable. If that is indeed the case, then HOMER could potentially be integrated into electronic health records as a decision-support tool. ■

**Author disclosures** are available with the text of this article at [www.atsjournals.org](http://www.atsjournals.org).

**Acknowledgment:** The authors thank Joseph Cicienia, M.D., Sonali Sethi, M.D., and Michael Machuzak, M.D., for their contribution to data collection.

## References

- Silvestri GA, Gonzalez AV, Jantz MA, Margolis ML, Gould MK, Tanoue LT, *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:e211S–e250S.
- National Comprehensive Cancer Network. Non-small cell lung cancer. Version 2.2019; 2018 [accessed 2019 Dec 13]. Available from: [https://www.nccn.org/professionals/physician\\_gls/pdf/nscl.pdf](https://www.nccn.org/professionals/physician_gls/pdf/nscl.pdf).
- Reck M, Rabe KF. Precision diagnosis and treatment for advanced non-small-cell lung cancer. *N Engl J Med* 2017;377:849–861.
- O'Connell OJ, Almeida FA, Simoff MJ, Yarmus L, Lazarus R, Young B, *et al*. A prediction model to help with the assessment of adenopathy in lung cancer: HAL. *Am J Respir Crit Care Med* 2017;195:1651–1660.
- Akthar AS, Ferguson MK, Koshy M, Vigneswaran WT, Malik R. Limitations of PET/CT in the detection of occult N1 metastasis in clinical stage I(T1-2aN0) non-small cell lung cancer for staging prior to stereotactic body radiotherapy. *Technol Cancer Res Treat* 2017;16:15–21.
- Vial MR, Khan KA, O'Connell O, Peng SA, Gomez DR, Chang JY, *et al*. Endobronchial ultrasound-guided transbronchial needle aspiration in the nodal staging of stereotactic ablative body radiotherapy patients. *Ann Thorac Surg* 2017;103:1600–1605.
- Hunink MGM, Weinstein MC, Wittenberg E, Drummond MF, Pliskin JS, Wong JB, *et al*. Choosing the best treatment. In: *Decision making in health and medicine: integrating evidence and values*, 2 ed. Cambridge: Cambridge University Press; 2014. pp. 53–77.
- Louie AV, Senan S, Patel P, Ferket BS, Lagerwaard FJ, Rodrigues GB, *et al*. When is a biopsy-proven diagnosis necessary before stereotactic ablative radiotherapy for lung cancer?: a decision analysis. *Chest* 2014;146:1021–1028.
- Martinez-Zayas G, Almeida FA, Simoff MJ, Yarmus L, Molina S, Young B, *et al*. Prediction model for malignant N1, N2 or N3 nodal disease relative to N0 in patients with non-small cell lung cancer [abstract]. *Am J Respir Crit Care Med* 2019;199:A1259.
- Ost DE, Ernst A, Lei X, Kovitz KL, Benzaquen S, Diaz-Mendoza J, *et al*.; AQuIRE Bronchoscopy Registry. Diagnostic yield and complications of bronchoscopy for peripheral lung lesions: results of the AQuIRE Registry. *Am J Respir Crit Care Med* 2016;193:68–77.
- Ost D, Simoff M, Feller-Kopman D, Kovitz K, Eapen G, Greenhill S, *et al*. Risk adjusted diagnostic yield and outcomes for endobronchial ultrasound guided TBNA: results of the ACCP AQuIRE Registry. *Chest* 2009;136:69S.
- Eapen GA, Shah AM, Lei X, Jimenez CA, Morice RC, Yarmus L, *et al*.; American College of Chest Physicians Quality Improvement Registry, Education, and Evaluation (AQuIRE) Participants.

Complications, consequences, and practice patterns of endobronchial ultrasound-guided transbronchial needle aspiration: results of the AQUIRE registry. *Chest* 2013;143:1044–1053.

13. Ost DE, Ernst A, Lei X, Feller-Kopman D, Eapen GA, Kovitz KL, et al.; AQUIRE Bronchoscopy Registry. Diagnostic yield of endobronchial ultrasound-guided transbronchial needle aspiration: results of the AQUIRE Bronchoscopy Registry. *Chest* 2011;140:1557–1566.
14. Dettlerbeck FC, Boffa DJ, Kim AW, Tanoue LT. The eighth edition lung cancer stage classification. *Chest* 2017;151:193–203.
15. Casal RF, Vial MR, Miller R, Mudambi L, Grosu HB, Eapen GA, et al. What exactly is a centrally located lung tumor? Results of an online survey. *Ann Am Thorac Soc* 2017;14:118–123.
16. Casal RF, Sepesi B, Sagar AS, Tschirren J, Chen M, Li L, et al. Centrally located lung cancer and risk of occult nodal disease: an objective evaluation of multiple definitions of tumour centrality with dedicated imaging software. *Eur Respir J* 2019;53:1802220.
17. Austin PC, van Klaveren D, Vergouwe Y, Nieboer D, Lee DS, Steyerberg EW. Validation of prediction models: examining temporal and geographic stability of baseline risk and estimated covariate effects. *Diagn Progn Res* 2017;1:12.
18. Steyerberg EW, Vergouwe Y. Towards better clinical prediction models: seven steps for development and an ABCD for validation. *Eur Heart J* 2014;35:1925–1931.
19. Steyerberg EW. Clinical prediction models: a practical approach to development, validation, and updating. New York, NY: Springer; 2009.
20. Rivera MP, Mehta AC, Wahidi MM. Establishing the diagnosis of lung cancer: diagnosis and management of lung cancer, 3rd ed: American College of Chest Physicians evidence-based clinical practice guidelines. *Chest* 2013;143:e142S–e165S.
21. Gould MK, Donington J, Lynch WR, Mazzone PJ, Midthun DE, Naidich DP, et al. Evaluation of individuals with pulmonary nodules: when is it lung cancer? Diagnosis and management of lung cancer, 3rd ed: American College of Chest Physicians evidence-based clinical practice guidelines. *Chest* 2013;143:e93S–e120S.
22. Shafazand S, Gould MK. A clinical prediction rule to estimate the probability of mediastinal metastasis in patients with non-small cell lung cancer. *J Thorac Oncol* 2006;1:953–959.
23. Zhang Y, Sun Y, Xiang J, Zhang Y, Hu H, Chen H. A prediction model for N2 disease in T1 non-small cell lung cancer. *J Thorac Cardiovasc Surg* 2012;144:1360–1364.
24. Chen K, Yang F, Jiang G, Li J, Wang J. Development and validation of a clinical prediction model for N2 lymph node metastasis in non-small cell lung cancer. *Ann Thorac Surg* 2013;96:1761–1768.
25. Koike T, Koike T, Yamato Y, Yoshiya K, Toyabe S. Predictive risk factors for mediastinal lymph node metastasis in clinical stage IA non-small-cell lung cancer patients. *J Thorac Oncol* 2012;7:1246–1251.
26. Farjah F, Lou F, Sima C, Rusch VW, Rizk NP. A prediction model for pathologic N2 disease in lung cancer patients with a negative mediastinum by positron emission tomography. *J Thorac Oncol* 2013;8:1170–1180.
27. Cho S, Song IH, Yang HC, Kim K, Jheon S. Predictive factors for node metastasis in patients with clinical stage I non-small cell lung cancer. *Ann Thorac Surg* 2013;96:239–245.
28. Silvestri GA, Gonzalez AV, Jantz MA, Margolis ML, Gould MK, Tanoue LT, 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:e211S–e250S.
29. Annema JT. Complete endosonographic intrathoracic nodal staging for lung cancer patients in whom SABR is considered. *ClinicalTrials.gov*; 2016 [accessed 2019 Dec 13]. Available from: <https://clinicaltrials.gov/ct2/show/NCT02997449>.
30. Wink KCJ, van Baardwijk A, Troost EGC, De Ruyscher D. Nodal recurrence after stereotactic body radiotherapy for early stage non-small cell lung cancer: incidence and proposed risk factors. *Cancer Treat Rev* 2017;56:8–15.
31. Ball D, Mai GT, Vinod S, Babington S, Ruben J, Kron T, et al.; TROG 09.02 CHISEL investigators. Stereotactic ablative radiotherapy versus standard radiotherapy in stage 1 non-small-cell lung cancer (TROG 09.02 CHISEL): a phase 3, open-label, randomised controlled trial. *Lancet Oncol* 2019;20:494–503.
32. Rowell NP, Williams CJ. Radical radiotherapy for stage I/II non-small cell lung cancer in patients not sufficiently fit for or declining surgery (medically inoperable). *Cochrane Database Syst Rev* 2001;(2):Cd002935.
33. Dosoretz DE, Katin MJ, Blitzer PH, Rubenstein JH, Salenius S, Rashid M, et al. Radiation therapy in the management of medically inoperable carcinoma of the lung: results and implications for future treatment strategies. *Int J Radiat Oncol Biol Phys* 1992;24:3–9.
34. Graham PH, GebSKI VJ, Langlands AO. Radical radiotherapy for early non-small cell lung cancer. *Int J Radiat Oncol Biol Phys* 1995;31:261–266.
35. Zhang HX, Yin WB, Zhang LJ, Yang ZY, Zhang ZX, Wang M, et al. Curative radiotherapy of early operable non-small cell lung cancer. *Radiother Oncol* 1989;14:89–94.
36. Jeremic B, Shibamoto Y, Acimovic L, Djuric L. Randomized trial of hyperfractionated radiation therapy with or without concurrent chemotherapy for stage III non-small-cell lung cancer. *J Clin Oncol* 1995;13:452–458.
37. Sibley GS, Jamieson TA, Marks LB, Anscher MS, Prosnitz LR. Radiotherapy alone for medically inoperable stage I non-small-cell lung cancer: the Duke experience. *Int J Radiat Oncol Biol Phys* 1998;40:149–154.
38. O'Rourke N, Roqué I Figuls M, Farré Bernadó N, Macbeth F. Concurrent chemoradiotherapy in non-small cell lung cancer. *Cochrane Database Syst Rev* 2010; (6):CD002140.
39. Blanke C, Ansari R, Mantravadi R, Gonin R, Tokars R, Fisher W, et al. Phase III trial of thoracic irradiation with or without cisplatin for locally advanced unresectable non-small-cell lung cancer: a Hoosier Oncology Group protocol. *J Clin Oncol* 1995;13:1425–1429.
40. Cakir S, Egehan I. A randomised clinical trial of radiotherapy plus cisplatin versus radiotherapy alone in stage III non-small cell lung cancer. *Lung Cancer* 2004;43:309–316.
41. Clamon G, Herndon J, Cooper R, Chang AY, Rosenman J, Green MR. Radiosensitization with carboplatin for patients with unresectable stage III non-small-cell lung cancer: a phase III trial of the Cancer and Leukemia Group B and the Eastern Cooperative Oncology Group. *J Clin Oncol* 1999;17:4–11.
42. Huber RM, Flentje M, Schmidt M, Pöllinger B, Gosse H, Willner J, et al.; Bronchial Carcinoma Therapy Group. Simultaneous chemoradiotherapy compared with radiotherapy alone after induction chemotherapy in inoperable stage IIIA or IIIB non-small-cell lung cancer: study CTRT99/97 by the Bronchial Carcinoma Therapy Group. *J Clin Oncol* 2006;24:4397–4404.
43. Jeremic B, Shibamoto Y, Acimovic L, Milisavljevic S. Hyperfractionated radiotherapy alone for clinical stage I non-small cell lung cancer. *Int J Radiat Oncol Biol Phys* 1997;38:521–525.
44. Jeremic B, Shibamoto Y, Acimovic L, Milisavljevic S. Hyperfractionated radiation therapy with or without concurrent low-dose daily carboplatin/etoposide for stage III non-small-cell lung cancer: a randomized study. *J Clin Oncol* 1996;14:1065–1070.
45. SchaaK-Koning C, van den Bogaert W, Dalesio O, Festen J, Hoogenhout J, van Houtte P, et al. Effects of concomitant cisplatin and radiotherapy on inoperable non-small-cell lung cancer. *N Engl J Med* 1992;326:524–530.
46. Soresi E, Clerici M, Grilli R, Borghini U, Zucali R, Leoni M, et al. A randomized clinical trial comparing radiation therapy v radiation therapy plus cis-dichlorodiammine platinum (II) in the treatment of locally advanced non-small cell lung cancer. *Semin Oncol* 1988; 15(Suppl 7):20–25.
47. Atagi S, Kawahara M, Tamura T, Noda K, Watanabe K, Yokoyama A, et al.; Japan Clinical Oncology Group (JCOG9812). Standard thoracic radiotherapy with or without concurrent daily low-dose carboplatin in elderly patients with locally advanced non-small cell lung cancer: a phase III trial of the Japan Clinical Oncology Group (JCOG9812). *Jpn J Clin Oncol* 2005;35:195–201.
48. Aupérin A, Le Péchoux C, Rolland E, Curran WJ, Furuse K, Fournel P, et al. Meta-analysis of concomitant versus sequential radiochemotherapy in locally advanced non-small-cell lung cancer. *J Clin Oncol* 2010;28:2181–2190.
49. Nyman J, Hallqvist A, Lund J-Å, Brustugun O-T, Bergman B, Bergström P, et al. SPACE - a randomized study of SBRT vs conventional fractionated radiotherapy in medically inoperable stage I NSCLC. *Radiother Oncol* 2016;121:1–8.

50. Ost DE, Gould MK. Decision making in patients with pulmonary nodules. *Am J Respir Crit Care Med* 2012;185:363–372.
51. Marchand C, Medford ARL. Relationship between endobronchial ultrasound-guided (EBUS)-transbronchial needle aspiration utility and computed tomography staging, node size at EBUS, and positron emission tomography scan node standard uptake values: a retrospective analysis. *Thorac Cancer* 2017;8:285–290.
52. Vial MR, O'Connell OJ, Grosu HB, Hernandez M, Noor L, Casal RF, 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;23:76–81.
53. Yasufuku K, Nakajima T, Waddell T, Keshavjee S, Yoshino I. Endobronchial ultrasound-guided transbronchial needle aspiration for differentiating N0 versus N1 lung cancer. *Ann Thorac Surg* 2013;96:1756–1760.
54. Leong TL, Loveland PM, Gorelik A, Irving L, Steinfors DP. Preoperative staging by EBUS in cN0/N1 lung cancer: systematic review and meta-analysis. *J Bronchology Interv Pulmonol* 2019;26:155–165.
55. Herth FJ, Eberhardt R, Krasnik M, Ernst A. Endobronchial ultrasound-guided transbronchial needle aspiration of lymph nodes in the radiologically and positron emission tomography-normal mediastinum in patients with lung cancer. *Chest* 2008;133:887–891.
56. Szlubowski A, Zielinski M, Soja J, Annema JT, Sosnicki W, Jakubiak M. A combined approach of endobronchial and endoscopic ultrasound-guided needle aspiration in the radiologically normal mediastinum in non-small-cell lung cancer staging—a prospective trial. *Eur J Cardiothorac Surg* 2010;37:1175–1179.
57. Sakairi Y, Hoshino H, Fujiwara T, Nakajima T, Yasufuku K, Yoshida S, et al. Validation of EBUS-TBNA-integrated nodal staging in potentially node-positive non-small cell lung cancer. *Gen Thorac Cardiovasc Surg* 2013;61:522–527.
58. Hwangbo B, Kim SK, Lee HS, Lee HS, Kim MS, Lee JM, et al. Application of endobronchial ultrasound-guided transbronchial needle aspiration following integrated PET/CT in mediastinal staging of potentially operable non-small cell lung cancer. *Chest* 2009;135:1280–1287.
59. Oki M, Saka H, Ando M, Kitagawa C, Kogure Y, Seki Y. Endoscopic ultrasound-guided fine needle aspiration and endobronchial ultrasound-guided transbronchial needle aspiration: are two better than one in mediastinal staging of non-small cell lung cancer? *J Thorac Cardiovasc Surg* 2014;148:1169–1177.
60. Shingyoji M, Nakajima T, Yoshino M, Yoshida Y, Ashinuma H, Itakura M, et al. Endobronchial ultrasonography for positron emission tomography and computed tomography-negative lymph node staging in non-small cell lung cancer. *Ann Thorac Surg* 2014;98:1762–1767.
61. Dooms C, Tournoy KG, Schuurbiens O, Decaluwe H, De Ryck F, Verhagen A, et al. Endosonography for mediastinal nodal staging of clinical N1 non-small cell lung cancer: a prospective multicenter study. *Chest* 2015;147:209–215.
62. Dong X, Qiu X, Liu Q, Jia J. Endobronchial ultrasound-guided transbronchial needle aspiration in the mediastinal staging of non-small cell lung cancer: a meta-analysis. *Ann Thorac Surg* 2013;96:1502–1507.
63. 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.
64. 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.
65. Evison M, Morris J, Martin J, Shah R, Barber PV, Booton R, et al. Nodal staging in lung cancer: a risk stratification model for lymph nodes classified as negative by EBUS-TBNA. *J Thorac Oncol* 2015;10:126–133.
66. Annema JT, van Meerbeeck JP, Rintoul RC, Dooms C, Deschepper E, Dekkers OM, et al. Mediastinoscopy vs endosonography for mediastinal nodal staging of lung cancer: a randomized trial. *JAMA* 2010;304:2245–2252.
67. Ernst A, Eberhardt R, Krasnik M, Herth FJ. Efficacy of endobronchial ultrasound-guided transbronchial needle aspiration of hilar lymph nodes for diagnosing and staging cancer. *J Thorac Oncol* 2009;4:947–950.
68. Jhun BW, Um S-W, Suh GY, Chung MP, Kim H, Kwon OJ, et al. Clinical value of endobronchial ultrasound findings for predicting nodal metastasis in patients with suspected lymphadenopathy: a prospective study. *J Korean Med Sci* 2014;29:1632–1638.
69. Yan J-H, Pan L, Chen X-L, Chen J-W, Yan L-M, Liu B, et al. Endobronchial ultrasound versus conventional transbronchial needle aspiration in the diagnosis of mediastinal lymphadenopathy: a meta-analysis. *Springerplus* 2016;5:1716.