Published in final edited form as:

Ann Thorac Surg. 2011 January ; 91(1): 227–233. doi:10.1016/j.athoracsur.2010.08.054.

Existing General Population Models Inaccurately Predict Lung Cancer Risk in Patients Referred for Surgical Evaluation

James M. Isbell, MD, MSCI^a, Stephen Deppen, MA, MS^b, Joe B. Putnam Jr, MD^{b,c,g}, Jonathan C. Nesbitt, MD^{b,c}, Eric S. Lambricht, MD^{b,c,g}, Aaron Dawes, BA^f, Pierre P. Massion, MD^{d,e,g}, Theodore Speroff, PhD^{e,g}, David R. Jones, MD^h, and Eric L. Grogan, MD, MPH^{b,c,g}

^aDepartment of Surgery, Nashville, TN

^bDepartment of Thoracic Surgery, Nashville, TN

^cVanderbilt-Ingram Cancer Center, Nashville, TN

^dDivision of Pulmonary and Critical Care Medicine, Nashville, TN

^eDepartment of Medicine, Nashville, TN

^fVanderbilt University School of Medicine, Nashville, TN

^gVeterans Affairs Tennessee Valley Healthcare System, Nashville, TN

^hDivision of Thoracic and Cardiovascular Surgery, University of Virginia, Charlottesville, VA

Abstract

Background—Patients undergoing resections for suspicious pulmonary lesions have a 9-55% benign rate. Validated prediction models exist to estimate the probability of malignancy in a general population and current practice guidelines recommend their use. We evaluated these models in a surgical population to determine the accuracy of existing models to predict benign or malignant disease.

Methods—We conducted a retrospective review of our thoracic surgery quality improvement database (2005-2008) to identify patients who underwent resection of a pulmonary lesion. Patients were stratified into subgroups based on age, smoking status and fluorodeoxyglucose positron emission tomography (PET) results. The probability of malignancy was calculated for each patient using the Mayo and SPN prediction models. Receiver operating characteristic (ROC) and calibration curves were used to measure model performance.

Results—89 patients met selection criteria; 73% were malignant. Patients with preoperative PET scans were divided into 4 subgroups based on age, smoking history and nodule PET avidity. Older smokers with PET-avid lesions had a 90% malignancy rate. Patients with PET- non-avid lesions, or PET-avid lesions with age<50 years or never smokers of any age had a 62% malignancy rate. The area under the ROC curve for the Mayo and SPN models was 0.79 and 0.80, respectively; however, the models were poorly calibrated ($p<0.001$).

Conclusions—Despite improvements in diagnostic and imaging techniques, current general population models do not accurately predict lung cancer among patients referred for surgical

Corresponding Author: Eric Grogan, M.D., M.P.H., Department of Thoracic Surgery, 609 Oxford House, Nashville, TN 37232 (phone) 615-322-006 (fax) 615-343-9194, eric.grogan@vanderbilt.edu.

Presented at the Forty-sixth Annual Meeting of The Society of Thoracic Surgeons, Ft. Lauderdale, Florida, January 25-27, 2010.

evaluation. Prediction models with greater accuracy are needed to identify patients with benign disease to reduce non-therapeutic resections.

Keywords

Lung Cancer; Lung Cancer; diagnosis; cancer staging; Positron Emission Tomography (PET)

Introduction

Over 150,000 patients in the United States are diagnosed with a focal pulmonary lesion annually [1]. With the growing use of computed tomography(CT), suspicious nodules are being detected with increasing frequency [2]. The prevalence of malignancy in these pulmonary nodules ranges from 10 to 68%, depending on patient selection, referral patterns and regional variation of granulomatous disease [1, 3-5]. Suspicious nodules must be promptly identified and treated given the relatively high malignancy rate found in these nodules, with the expectation of improved survival in patients with early stage cancers [1, 3].

Current guidelines recommend that clinicians estimate that the likelihood a focal pulmonary lesion is malignant using either best clinical judgment or a validated prediction model [6]. Ideally, the clinical prediction tool accurately discriminates between benign and malignant nodules in patients referred for resection. This tool should improve patient outcomes by providing resection for patients with malignant nodules, and deferring resection with its attendant risks for patients with benign disease. Several clinical prediction models have been developed to calculate a pretest probability of malignancy in a general population to guide clinicians in the interpretation of subsequent diagnostic test results [7-13]. However, these prediction tools may not apply to patients seen by thoracic surgeons due to differences in the prevalence of cancer and the spectrum of disease.

Even combining expert clinical judgment, predictive models, computed tomography and fluorodeoxyglucose positron emission tomography (PET), 9-55% of patients undergoing resection have benign disease [4, 5, 14-17]. To determine the accuracy of existing predictive models (i.e., the Mayo Model and the SPN Model) and understand which factors identify patients more likely to receive non-therapeutic resections, we evaluated patients following resection of a focal pulmonary lesion. We sought to: 1) assess patients classified into groups based on age, smoking history, and PET avidity with focal pulmonary lesions for likelihood of benign or malignant disease and 2) evaluate the accuracy of two existing lung cancer prediction models for the surgical patient population.

Methods

Population

We identified eligible patients who underwent an operation for a focal pulmonary lesion from January 1, 2005 to December 31, 2008 using Vanderbilt University Medical Center's Thoracic Surgery Quality Improvement database. All patients had a final postoperative pathologic diagnosis; those with known metastatic lung cancer prior to surgery were excluded. 190 patients were reviewed. One nodule diameter could not be found and was excluded from analysis leaving 189 patients in the final dataset. The Vanderbilt University Institutional Review Board approved this study and waived the need to obtain individual patient consent (IRB #081298).

Data Preparation and Classification

Additional imaging data were collected using retrospective chart review. Imaging variables included maximum diameter as reported by radiologists from CT, magnetic resonance imaging (MRI) or chest radiograph; lesion growth observed in any serial imaging; lesion shape characterized as smooth, lobulated, or spiculated; PET avidity and lesion location.

PET avidity was recorded as 'positive' if a standard uptake value (SUV) of 2.5 or higher was reported. When no SUV was reported, any indication in the radiologist's report of malignancy, such as "cancerous", "likely", "possible", or "probable" were classified as avid. Cases that were ambiguous by radiologist report were reviewed and classified by one of two surgeon authors (ELG or JMI), who were blinded to the malignancy status. Lesion edge characteristics were defined by the terms "smooth", "lobulated", "lobular", "lobed", "coronal", "corona", "spikey" or "spiculated" in the radiologist's report. If one of these terms was not specified in the report, the lesion edge characteristic was designated missing for SPN Model probability and "not spiculated" for estimating the Mayo Clinic Model probability of malignancy.

The patients were classified into four subgroups based upon characteristics clinically significant to the surgeon: age, smoking history and nodule PET avidity (Table 2). More specifically, Group 1 consisted of those patients ≥ 50 years of age + ever smokers (i.e., current or former) + PET avid lesions. Group 2 patients were ≥ 50 years of age + ever smokers + PET non-avid lesions. Group 3 patients were <50 years of age or never smokers + PET avid lesions. Group 4 patients were <50 years of age or never smokers + PET non-avid lesions. Group 1 was considered the high risk subgroup and used as the reference group. Groups 2, 3, and 4 were considered lower risk and compared to Group 1.

Prediction Models

The predicted probability of malignancy was calculated using two previously published models: the Mayo Clinic Model [7] and the SPN Model [18].

Mayo Clinic Model—The Mayo Clinic Model is a general population model and developed to aid the general practitioner in determining risk of lung cancer in an individual patient. This model is defined by the equation: $\text{pre-test probability of malignancy} = e^x / (1 + e^x)$, where $x = -6.8272 + (0.0391 * \text{age in years}) + (0.7917 * \text{smoking status}) + (1.3388 * \text{previous cancer}) + (0.1274 * \text{nodule diameter}) + 1.0407 * \text{spiculated edge} + (0.7838 * \text{upper lobe})$. The Mayo Clinic Model criteria exclude all who had a previous thoracic cancer or a non-thoracic cancer within the last five years and includes only those lesions between 4 and 30 mm in diameter.

Solitary Pulmonary Nodule Prediction Model (SPN Model)—The SPN Model follows the Bayesian method of Gurney and Dewan [10-12]. The Bayesian model uses previously published likelihood ratios for malignancy to adjust the conditional probability of malignancy. The initial prior probability can be a subjective clinical estimate or based on the prevalence in historical patient populations. We used 73%, our population malignancy rate, and 50%, an uninformed prior probability, as the prior probabilities of cancer. The SPN Model includes age (in years), pack years of cigarette smoking, hemoptysis, previous non-thoracic malignancy, nodule diameter, nodule location, edge characteristics, growth rate, cavity wall thickness, calcification, CT contrast enhancement results, and PET results. Calcification and CT contrast enhancement results were reported on only 6 individuals in our dataset and were not used in our analysis.

Data Analysis

Differences in clinical and radiological characteristics between patients with benign and malignant lesions were tested using the Student t-test for means, Mann-Whitney U test for medians, and the binomial test or Chi-square tests for proportions. The level of statistical significance was set to the 1% level ($P < 0.01$) to adjust for multiple comparisons where applicable.

ROC and Calibration Curve Analysis—Receiver operating curves (ROC) and areas under the curve measured discrimination of model performance. To perform model calibration analyses, our surgical population was rank ordered and divided into 5-equally weighted groups (quintiles) based upon the probability of malignancy as predicted by the Mayo or SPN models. Model calibration was evaluated by comparing observed to predicted probability of cancer among the quintiles and tested for statistical significance using the Hosmer-Lemeshow goodness of fit test in each model. We then compared the accuracy of the predicted Mayo and SPN models with the actual rate of malignancy in Groups 1 - 4 to determine if the variables clinically significant to the surgeon (age, smoking status, PET avidity) affected model fit. The analysis was repeated for patients who did not have a pre-operative pathological diagnosis of lung cancer and in patients who only fit the original Mayo Model inclusion criteria. Data were analyzed using Stata 10.0 (College Station, TX).

Results

The overall prevalence of malignancy was 73% (Table 1). Patients with lung cancer were older, had a history of smoking, and smoked more even among those who did smoke, exhibited lower FEV₁, had larger lesion diameters, and were more likely to have PET avid lesions. Seventy-seven patients had a pathological diagnosis prior to their resection, and three of those were incorrect diagnoses. PET results were available for 145 patients. Ninety-nine of the 115 malignant lesions (86%) were PET avid, and 19 of 30 benign lesions (63%) were PET avid. Spiculation was observed in 48 of 189 patients through CT or chest radiograph. Forty-one of those 48 patients had cancer (85%) and 7 of 48 (15%) had benign disease.

Identification of Populations likely to have Benign Resections

The observed prevalence of lung cancer was determined in Groups 1-4 based upon characteristics clinically significant to the surgeon: age, smoking status, and PET avidity (Table 2). Each of the 3 lower risk subgroups was compared to Group 1 (subgroup with a 90% malignancy rate: older smokers with PET avid lesions). This compares to 65% for Group 2 ($p = 0.006$), 67% for Group 3 ($p = 0.004$), and 46% for Group 4 ($p < 0.001$). In aggregate, Groups 2, 3 and 4 malignancy rate (62%) was significantly lower than the reference group of older smokers with FDG-PET avid lesions ($p < 0.001$). Patients with no preoperative tissue diagnosis ($n = 112$) had a 57% malignancy rate. The 92 patients who met all original Mayo Model inclusion criteria had an observed cancer rate of 65%. The corresponding mean predicted cancer rates for the Mayo and SPN models were calculated for all patients and for each of the clinical subgroups (Table 2).

Mayo and SPN Model Analyses

The areas under the ROC curve (AUC) for the Mayo and SPN models were 0.78 (95%CI 0.70 to 0.85) (Figure 1A) and 0.80 (95%CI 0.73 to 0.87) (Figure 2A), respectively. The probability of cancer was estimated with the SPN Model using the observed prevalence of 73% as the prior probability. Of the 189 patients, 92 met the original Mayo Model inclusion criteria. Among these 92 patients, the AUC values for the Mayo and SPN models were 0.84 (95%CI 0.75 to 0.92) and 0.81 (95%CI 0.72 to 0.90), respectively. Because limiting the

analysis to patients who met the strict inclusion criteria of the Mayo model had only a minimal effect on the AUC values, we included all 189 patients in the remainder of our analyses. Among older smokers with PET avid lesions, the Mayo and SPN model AUCs were 0.64 (95%CI 0.45 to 0.82) and 0.70 (95%CI 0.52 to 0.89), respectively. The Group 2, 3 and 4 patients (those designated with a higher likelihood of benign disease) had AUCs of 0.83 (95%CI 0.71 to 0.94) and 0.76 (95%CI 0.63 to 0.94) for the Mayo and SPN models, respectively.

Among all patients referred to a thoracic surgeon for evaluation, the observed malignancy prevalence was 73% compared to the Mayo and SPN models' predicted probabilities of 60% (95%CI 55 to 65%) and 71% (95%CI 65 to 76%), respectively (Table 2). The SPN Model probability of malignancy was also calculated for each individual using an uninformed, 50% prior probability of cancer. In all analysis of the accuracy of the SPN Model, the lower 50% prior probability reduced the average predicted probability of cancer in each population by 2-3%. The calibration data for the Mayo and SPN models are shown in Figures 1B and 2B, respectively. Goodness-of-fit tests demonstrate that the Mayo and SPN models did not perform well when used to predict risk of malignancy in the surgical patient population ($p < 0.001$). The Mayo model was well calibrated for the two highest quintiles but underestimated the probability of malignancy for the lower quintiles. The SPN Model underestimated the probability of malignancy for the lower quintiles and overestimated in the higher quintiles.

We evaluated the predictive ability of each model based upon characteristics clinically relevant to surgeons (Table 2). The Mayo Model had a predicted probability of only 72% in Group 1, significantly underestimating the prevalence of malignancy in this high risk group. The Mayo Model also underestimated the probability of malignancy for Groups 2-4. The SPN Model performed well in Group 1 patients (i.e., older smokers with PET avid nodules) with observed prevalence and predicted probability of malignancy being 90% and 92%, respectively and Group 3 (67% observed versus 73% predicted). Malignancy prevalence was underestimated in Groups 2 and 4 by the SPN Model. When we excluded all patients who did not fit the strict Mayo Model criteria, we found an even greater discrepancy between observed (65%) and predicted (39%) malignancy prevalence (Table 2).

Comment

In this population of patients undergoing surgical resection for known or suspected lung cancer, the overall malignancy rate was 73%. We identified subpopulations with varying prevalence of cancer. Patients who are older smokers with PET avid lesions had a 90% malignancy rate. All other patients have a higher likelihood of benign disease but still had a 62% malignancy rate. Those patients without a preoperative tissue diagnosis had a 57% malignancy rate.

In our cohort of patients who underwent resection for a focal pulmonary lesion, we used two existing clinical prediction tools to estimate the probability of malignancy. Both performed poorly and were not calibrated to the surgical population. The Mayo model, which has been previously validated [7, 13, 19], underestimated the probability of malignancy in our patient population. The SPN model, while more accurate than the Mayo model overall, significantly underestimated the observed frequency of malignancy in the lowest risk subgroup of patients with age less than 50 or never smokers (SPN calculated 8% vs. observed 45%).

Current practice guidelines recommend using prediction models to estimate the pretest probability of malignancy [6]. Application of these models may be most valuable in a general practice population, but less useful in patients being evaluated for resection by a

thoracic surgeon. Clinicians must choose a likelihood threshold where additional invasive/ diagnostic procedures or surgical resection would be warranted. Factors usually considered in setting the probability threshold are the clinician's level of suspicion for cancer, the overall health status of the patient and whether the patient could tolerate surgical resection. To improve the chance of treating a suspected cancer, surgeons generally set a threshold that reflects a high sensitivity for cancer, at the expense of decreased specificity. In a physically fit patient, resection may be easily tolerated; however, in less fit patients, the risks of the proposed surgical treatment may overshadow the benefits - particularly for those patients with benign disease.

Existing prediction models were not designed or validated using a surgical patient population. Non-surgical clinicians often observe patients with pulmonary nodules using a watch-and-wait strategy, referring them to a surgeon only after growth has been detected or if they are PET avid. Accordingly, these patients referred for surgical evaluation have a higher prevalence of cancer. Unless the cancer prevalence of the population being evaluated is similar to that used to develop the model, any model based upon a population of lesser risk of cancer will not represent the true distribution of risk and greatly under- or over-estimate the individual probability of cancer in the individual or in subgroups.

Two recent studies further support our results in different patient populations. Schultz et al [19] performed a validation of the Mayo model using data from 147 patients with pulmonary nodules in a population undergoing radiological evaluation. The prevalence of malignancy in this group was 45% with an AUC similar to ours of 0.78 (95%CI 0.71 to 0.86). This analysis demonstrated a minimal effect on model accuracy when patients with prior cancers were included. Herder et al [13] calculated an AUC of 0.79 (95%CI 0.70 to 0.87) using data from a radiological cohort of 106 patients with a 57% malignancy rate. Both of these studies found that the Mayo model underestimated the overall probability of cancer.

In our study, information on nodule size, location, spiculation, calcification and PET avidity was obtained from radiology reports. Over 60% of our patients were referred from outside our institution. Standardized uptake values and quantitative levels of contrast enhancement were not available for most of our patients. The variability within radiology reports and inconsistent availability of imaging files decreases the accuracy of developing any imaging-based predictive model. Structured reporting of radiographic findings and picture archive would facilitate future model development. We included only patients who underwent resection of their pulmonary nodules. This selection bias predisposes to those patients most likely to have a malignancy. If we included those patients who were evaluated by the surgeon and did not undergo a resection, then we may increase the specificity of our results, increase the observed AUC, and decrease the sensitivity for any cancers missed. Third, we included patients who did not conform to the Mayo Model strict selection criteria. Specifically, we included patients with pulmonary lesions greater than 3 cm, patients with prior history of lung cancer and those with a history of extrathoracic cancer diagnosed within 5 years of nodule identification. Including these groups expands the clinical applicability of the model. Through sub-analyses we demonstrated that each of these groups of patients had minimal effect on accuracy and calibration when included in the Mayo model. In addition to using the 73% prior probability of disease, we also used an uninformed, 50% prior probability of disease in the SPN model to estimate the probability of lung cancer in our analysis. With the uninformed prior probability, the mean predicted rate of cancer in each population was 2-3% less than that of the overall population and did not materially change conclusions or results compared to using the overall prevalence for the prior probability.

In summary, we show that available clinical prediction models for a general population perform relatively poorly in a surgical population. For an individual patient, this could lead to missed cancers and unnecessary additional diagnostic tests and nontherapeutic pulmonary resections. Accordingly, surgeons should use existing prediction tools to make operative decisions for patients referred for surgical evaluation with caution. Our study also identifies a surgical population with a higher likelihood of benign disease who would commonly undergo resection. This sub-population would benefit most from a more accurate prediction model. Development of future clinical prediction models for thoracic surgery must be designed and validated for the surgical patient population. Applying such a surgery-focused clinical predictive model to patients with a higher likelihood of benign disease would potentially reduce the number of pulmonary resections for benign disease and decrease the healthcare costs associated with the diagnosis and management of the focal pulmonary lesion.

Acknowledgments

This research was supported by: Vanderbilt Physician Scientist Development Award (E.L.G.), SPECS in lung cancer U01 CA114771 (P.P.M.) the lung SPORE CA90949 (P.P.M.), and a Merit award from the Department of Veterans Affairs (P.P.M.).

References

1. Tan B, Flaherty K, Kazerooni E, Iannettoni M. The solitary pulmonary nodule. *Chest*. 2003; 123:89S–96S. [PubMed: 12527568]
2. Bach PB, Silvestri GA, Hanger M, Jett JR. Screening for Lung Cancer ACCP guidelines *Chest*. 2007; 132:69S–77S.
3. Wahidi MM, Govert JA, Goudar RK, Gould MK, McCrory DC. Evidence for the treatment of patients with pulmonary nodules: when is it lung cancer?: ACCP evidence-based clinical practice guidelines. *Chest* (2nd edition). 2007; 132(3 Suppl):94S–107S. [PubMed: 17873163]
4. Grogan EL, Stukenborg GJ, Nagji AS, et al. Radiotracer-guided thoracoscopic resection is a cost-effective technique for the evaluation of subcentimeter pulmonary nodules. *Ann Thorac Surg*. 2008; 86(3):934–40. discussion 934–40. [PubMed: 18721586]
5. Smith MA, Battafarano RJ, Meyers BF, et al. Prevalence of Benign Disease in Patients Undergoing Resection for Suspected Lung Cancer. *The Annals of Thoracic Surgery*. 2006; 81(5):1824–1829. [PubMed: 16631679]
6. Gould MK, Fletcher J, Iannettoni MD, et al. Evaluation of patients with pulmonary nodules: when is it lung cancer?: ACCP evidence-based clinical practice guidelines. *Chest* (2nd edition). 2007; 132(3 Suppl):108S–130S. [PubMed: 17873164]
7. Swensen S, Silverstein M, Ilstrup D, Schleck C, Edell E. The probability of malignancy in solitary pulmonary nodules. Application to small radiologically indeterminate nodules. *Arch Intern Med*. 1997; 157:849–855. [PubMed: 9129544]
8. Spitz MR, Etzel CJ, Dong Q, et al. An Expanded Risk Prediction Model for Lung Cancer. *Cancer Prev Res*. 2008; 1(4):250–254.
9. Gould MK, Ananth L, Barnett PG. A clinical model to estimate the pretest probability of lung cancer in patients with solitary pulmonary nodules. *Chest*. 2007; 131(2):383–8. [PubMed: 17296637]
10. Gurney JW. Determining the likelihood of malignancy in solitary pulmonary nodules with Bayesian analysis. Part I. Theory. *Radiology*. 1993; 186(2):405–413. [PubMed: 8421743]
11. Gurney JW, Lyddon DM, McKay JA. Determining the likelihood of malignancy in solitary pulmonary nodules with Bayesian analysis. Part II. Application. *Radiology*. 1993; 186(2):415–422. [PubMed: 8421744]
12. Dewan NA, Shehan CJ, Reeb SD, et al. Likelihood of malignancy in a solitary pulmonary nodule: comparison of Bayesian analysis and results of FDG-PET scan. *Chest*. 1997; 112(2):416–22. [PubMed: 9266877]

13. Herder GJ, van Tinteren H, Golding RP, et al. Clinical Prediction Model To Characterize Pulmonary Nodules. *Chest*. 2005; 128(4):2490–2496. [PubMed: 16236914]
14. Veronesi G, Bellomi M, Scanagatta P, et al. Difficulties encountered managing nodules detected during a computed tomography lung cancer screening program. *J Thorac Cardiovasc Surg*. 2008; 136(3):611–7. [PubMed: 18805261]
15. Davies B, Ghosh S, Hopkinson D, Vaughan R, Rocco G. Solitary pulmonary nodules: pathological outcome of 150 consecutively resected lesions. *Interact Cardiovasc Thorac Surg*. 2005; 4(1):18–20. [PubMed: 17670346]
16. Stiles BM, Altes TA, Jones DR, et al. Clinical experience with radiotracer-guided thorascopic biopsy of small, indeterminate lung nodules. *Ann Thorac Surg*. 2006; 82(4):1191–6. discussion 1196–7. [PubMed: 16996906]
17. Reed C, Harpole D, Posther K, et al. Results of the American College of Surgeons Oncology Group Z0050 trial: the utility of positron emission tomography in staging potentially operable non-small cell lung cancer. *J Thorac Cardiovasc Surg*. 2003; 126:1943–1951. [PubMed: 14688710]
18. Gurney, JW. Chest x-ray SPN risk estimator. 2009. Available from: <http://www.chestx-ray.com/SPN/SPNProb.html>
19. Schultz EM, Sanders GD, Trotter PR, et al. Validation of two models to estimate the probability of malignancy in patients with solitary pulmonary nodules. *Thorax*. 2008; 63(4):335–41. [PubMed: 17965070]

Figure 1.

(A) ROC curve for Mayo Clinic model. Area under the ROC curve (AUC) = 0.78 (95% CI 0.70 to 0.85). (B) The calibration curve plots median predicted probability of a malignant nodule by observed frequency for patients in each of the 5 quintiles of predicted probability. A point above the diagonal indicates that the model underestimates the likelihood of cancer. A point below the diagonal indicates that the model overestimates the likelihood of cancer. A point near the diagonal indicates that the model is calibrated and the estimated probability is close to the observed probability of cancer. Goodness of fit estimated by Hosmer-Lemeshow (H-L) test ($p < 0.001$). P values > 0.05 are consistent with models that fit observed data.

Figure 2.

(A) ROC curve SPN model. Area under the ROC curve (AUC) = 0.80 (95% CI 0.73 to 0.87).

(B) The calibration curve plots median predicted probability of a malignant nodule by observed frequency for patients in each quintile of predicted probability. Goodness of fit estimated by Hosmer-Lemeshow test (H-L test $p < 0.001$). P-values > 0.05 are consistent with models that fit observed data.

Table 1

Clinical Characteristics of Patients

Characteristic	Overall (n = 189)	Malignant (n = 138)	Benign (n = 51)	p value
Age - median (Range)	63 (54,72)	65 (58,74)	53 (42,65)	<.001 ^a
Male (%)	95 (50)	73 (53)	22 (43)	0.14
Smoking history				
Current smoker (%)	40 (21)	33 (24)	7 (14)	0.13
Former smoker (%)	100 (53)	80 (58)	21 (41)	0.02
Never smoker (%)	49 (26)	25 (18)	23 (45)	0.001
Mean pack-years (SD)	42 (26)	45 (26)	30 (26)	0.001
Preop diagnosis of NSCLC	77 (41)	74 (53)	3 (6)	
History of prior cancer	49 (27)	37 (27)	12 (24)	0.43
Mean FEV ₁ (liters) ^b (SD) (n=160)	2.4 (0.8)	2.3 (0.8)	2.7 (0.9)	0.001
Mean Zubrod performance	0.53	0.53	0.54	.82
Lesion Size – median (range)	24 (15, 40)	27 (18, 45)	17 (10, 27)	0.001 ^a
AJCC Stage ^c				
I		71 (51)		
II		13 (9)		
III		32 (23)		
IV		5 (4)		
Other Cancer/Limited Small Cell		17 (13)		

Data presented as number (%) unless otherwise specified.

SD=standard deviation; NSCLC=non-small cell lung cancer; FEV₁=forced expiratory volume

^aMann-Whitney U test.

^bIndicates variable was not available for all patients, n=number of cases for which data were available. Missing data were excluded from statistical testing.

^cAmerican Joint Committee on Cancer (AJCC) 6th edition.

Table 2

Observed and Predicted Malignancy Prevalence

	Patients	Observed Cancer Prevalence	<i>p</i> value ^a	Predicted Cancer Prevalence	
				Mayo	SPN
Overall	189	73%		60%	71%
PET imaging	145	79%		64%	77%
Age ≥ 50 years + Ever Smokers	107	86%			
PET Avid (Group 1)	90	90%	Ref	72%	92%
PET Non-Avid (Group 2)	17	65%	0.006	61%	46%
Age < 50 years or Never Smokers	38	60%			
PET Avid (Group 3)	27	67%	0.004	51%	73%
PET Non-Avid (Group 4)	11	45%	<0.001	36%	8%
Groups 2-4 - Higher Likelihood of Benign Disease	55	62%	<0.001	51%	52%
No prior tissue diagnosis	112	57%		51%	65%
Original Mayo inclusion criteria ^b	92	65%		39%	62%

^a *p* values were obtained by comparing Groups 2-4 to Group 1 (Ref), univariate t-test.

^b Only patients with nodule diameter 4-30mm, no extrathoracic cancers within 5 years of nodule identification and no prior history of lung cancer were included in these subgroups.