The clinical stage of non-small cell lung cancer as assessed by means of fluorodeoxyglucose–positron emission tomographic/computed tomographic scanning is less accurate in cigarette smokers

Ayesha S. Bryant, MSPH, MD,^a and Robert James Cerfolio, MD, FACS, FCCP^b

Objective: The treatment of non-small cell lung cancer depends on the stage, and this is clinically best determined by using fluorodeoxyglucose–positron emission tomography/computed tomography. We evaluated the effect smoking has on the accuracy of this test.

Methods: We performed a prospective cohort study evaluating the accuracy of clinical stage compared with pathologic stage between cigarette smokers and nonsmokers with non-small cell lung cancer. All patients were assigned a clinical TNM stage after fluorodeoxyglucose–positron emission tomographic/computed tomographic scanning and then underwent meticulously pathologic TNM staging. If N2, N3, or M1 negative, patients underwent thoracotomy with complete thoracic lymphadenectomy. The clinical and pathologic stages were compared.

Results: There were 246 patients: 52 never smoked (NS group), 112 quit at least 1 month before fluorodeoxyglucose–positron emission tomography/computed tomography (Q group), and 82 were still smokers (S group). The 3 groups were similar for stage and histology. The overall accuracy was 83%, 80%, and 64% for the NS, Q, and S groups, respectively ($P = .03$). The accuracy for the T status was 88%, 84%, and 86%; accuracy for the N2 lymph nodes was 96%, 75%, and 72%; and accuracy for the N1 lymph nodes was 92%, 78%, and 80%, respectively, favoring the NS group. The greater the pack-year history, the greater the N2 inaccuracy ($P = .04$). Multivariate analysis showed that status of smoking ($P = .026$) and maxSUV value ($P = .014$) were independent predictors of fluorodeoxyglucose–positron emission tomography/computed tomography accuracy.

Conclusions: Patients with non-small cell lung cancer who continue to smoke at the time of their fluorodeoxyglucose–positron emission tomographic/computed tomographic scan have less accurate clinical staging compared with those who stopped 1 month before or who never smoked. As the pack-years increase, the accuracy for the N2 nodes decrease. Nonsmokers have the most accurate clinical staging.

From the Department of Epidemiology, School of Public Health,^a and the Division of Cardio-Thoracic Surgery, Department of Surgery,^b University of Alabama at Birmingham (UAB), Birmingham, Ala.

Read at the Eighty-sixth Annual Meeting of The American Association for Thoracic Surgery, Philadelphia, Pa, April 29-May 3, 2006.

Received for publication May 8, 2006; revisions received June 28, 2006; accepted for publication July 12, 2006.

Address for reprints: Robert J. Cerfolio, MD, Division of Cardiothoracic Surgery, University of Alabama at Birmingham, 1900 University Blvd, THT 712, Birmingham, AL 35294 (E-mail: Robert.cerfolio@ccc.uab.edu).

J Thorac Cardiovasc Surg 2006;132:1363-8
0022-5223/\$32.00

Copyright © 2006 by The American Association for Thoracic Surgery

doi:10.1016/j.jtcvs.2006.07.032

The treatment of non-small cell lung cancer (NSCLC) depends on the clinical stage that is often generated based on the result of an integrated fluorodeoxyglucose–positron emission tomographic/computed tomographic scan (FDG-PET/CT).¹ A dedicated positron emission tomographic (PET) scan now probably represents the new standard of care for noninvasive staging in patients with suspected or biopsy-proved NSCLC. We and others have shown that even though integrated FDG-PET/CT is the most accurate clinical staging tool currently available, it often is wrong when compared with the actual pathologic stage.²⁻⁵

The purpose of this study was to determine whether the staging accuracy of integrated FDG-PET/CT scanning was affected by cigarette smoking status. Thus

Abbreviations and Acronyms

CT	= computed tomography
FDG	= fluorodeoxyglucose
FDG-PET/CT	= fluorodeoxyglucose–positron emission tomographic/computed tomographic
MaxSUV	= maximum standardized uptake value
NSCLC	= non–small cell lung cancer
PET	= positron emission tomography
ROI	= region of interest

we evaluated accuracy in patients who never smoked cigarettes, those who quit at least 1 month before their PET scan, and those who were still smoking at the time of their PET scan. In addition, we also sought to evaluate whether a dose-response relationship exists between accuracy and pack-years smoked.

Methods**Patient Selection**

This is a prospective cohort study. Patients who presented to one thoracic surgeon over a 2-year period with an indeterminate pulmonary nodule or biopsy-proved NSCLC who underwent integrated FDG-PET/CT at the University of Alabama at Birmingham center were eligible for this study. Patients were excluded if they were less than 19 years old, had a history of type I diabetes, and/or received preoperative chemotherapy or radiation. The University of Alabama at Birmingham's institutional review board approved both the electronic prospective database used for this study and this prospective trial. Patient consent was obtained to enter their

data in the prospective database but not for entry into this specific study.

Radiologic Imaging

FDG-PET/CT scans were performed on an integrated PET/CT scanner (GE Discovery LS PET-CT Scanner; GE, Milwaukee, Wis). Patients were asked to fast for 4 hours and then subsequently received 555 MBq (15 mCi) of FDG intravenously, followed by PET after 1 hour. The scans were performed from the skull base to the midthigh level. The computed tomographic (CT) examination was used for attenuation correction of PET images. The scanning time for emission PET scanning was 5 minutes per bed position. Iterative reconstruction with CT attenuation correction was performed. The most recent CT scan of the chest was also available for visual correlation. Maximum standardized uptake value (maxSUV) of the primary tumor and of each suspicious lymph node station was determined by drawing regions of interest (ROIs) on the attenuation-corrected FDG-PET images around it. It was then calculated by using the software contained within the PET or PET/CT scanner with the following formula⁶: $MaxSUV = (C(\mu Ci/mL) / ID(\mu Ci)) / w(kg)$, where C is defined as activity at a pixel within the tissue defined by an ROI, and ID is defined as the injected dose per kilogram of the patient's body weight (w). The maxSUV within the selected ROIs was used throughout this study exclusively.

Procedures, Staging, and Surgical Intervention

All patients were clinically staged with the TNM classification system.⁷ A clinical stage was assigned for the patient based on a combination of CT scan results, FDG-PET/CT scan results, and the maxSUV of the primary tumor, nodes, or both by one physician (RJC). The size of the tumor was calculated by multiplying the greatest reported length and width of the lesion. (Because the

TABLE 1. Patient characteristics

Characteristic	Never smoked, n = 52	Former smokers, n = 112	Current smokers, n = 82	P value*
Median age (y)	55.3 ± 8.2	67.6 ± 5.2	64.4 ± 9.3	.50
Sex				
Male	20 (38%)	58 (52%)	64 (78%)	.002
Female	32 (62%)	54 (48%)	18 (22%)	
Race				
White	48 (92%)	104 (93%)	77 (94%)	.90
Black	4 (8%)	8 (7%)	5 (6%)	
Mean ECOG score	0.1	0.2	0.5	.78
Median packs per day (PPD)	0	1.4	1.5	Not applicable
Median years smoked (YS)	0	40	40	
Packs per year (PPD × YS)	0	56	60	
Quit (mo)	0	134	0	
Preoperative pulmonary function (median)				
FEV ₁ %	100%	75%	71%	.016
DLCO%	86%	67%	68%	.037
MVV%	85%	64%	71%	.110

ECOG, Eastern Cooperative Oncology Group; FEV₁%, percentage forced expiratory volume in 1 second; DLCO%, diffusing capacity of lung for carbon monoxide; MVV%, maximum ventilation volume. *Compares smokers with those who never smoked.

TABLE 2. Tumor characteristics

Characteristic	Never smoked, n = 52	Former smokers, n = 112	Smokers, n = 82	P value
Median maxSUV of tumor	7.3	5.1	4.9	.026
Median maxSUV of N2 lymph nodes that were ≥ 2.5	5.6	4.2	4.3	.09
Median size (length $\times$ width) cm ²	2 $\times$ 1.8 3.6 cm ²	1.5 $\times$ 2.0 3.0 cm ²	1.6 $\times$ 2.1 3.36 cm ²	.923
Histology				
Squamous cell	16 (31%)	53 (48%)	30 (37%)	.109*
Adenocarcinoma	18 (35%)	45 (40%)	37 (45%)	
Carcinoid	8 (14%)	3 (3%)	6 (7%)	
Bronchoalveolar	4 (6%)	2 (2%)	3 (3%)	
Other	6 (11%)	9 (8%)	6 (7%)	
Pathologic stage				
I	20 (38%)	70 (61%)	47 (57%)	.186
II	9 (17%)	16 (14%)	11 (13%)	
III	18 (35%)	22 (20%)	20 (24%)	
T3 N1 M0	3	3	3	
T1-T3 N2 M0	11	16	13	
T1-T3 N3 M0	3	2	2	
T4 Nx M0	1	1	2	
IV	5 (10%)	4 (4%)	4 (5%)	
Differentiation				
Well	8 (38%)	14 (23%)	12 (20%)	.192
Moderate	7 (33%)	22 (37%)	32 (52%)	
Poor	6 (29%)	24 (40%)	17 (28%)	
Lymphovascular invasion	6/20 (30%)	23/47 (49%)	13/29 (45%)	.295

maxSUV, maximum standardized uptake value. *Compares squamous cell and adenocarcinoma with carcinoid and bronchoalveolar carcinoma.

thickness was rarely reported, we report tumor size as square centimeters instead of cubic centimeters.) All suspicious N2, N3, or M1 areas (maxSUV >2.5) determined either by means of FDG-PET/CT scans, CT scans, or both underwent biopsy before pulmonary resection. Mediastinoscopy was used for biopsy of suspicious lymph nodes in the paratracheal area (stations 2R, 4R, 2L, and 4L and the superior part of the 7) and/or endoscopic transesophageal ultrasonographic fine-needle aspiration was used for biopsy of suspicious posterior aortopulmonary window (n = 5), subcarinal (n = 7), periesophageal (n = 8) and inferior pulmonary ligament (n = 9) lymph nodes. Endoscopic ultrasonography was performed during conscious sedation, as previously described.^{8,9}

Patients with suspected M1 disease in the liver, adrenal glands, or contralateral lung underwent definitive biopsy to prove or disprove M1 cancer. If the bone or brain was suspected to harbor metastases, magnetic resonance imaging was considered the standard reference. If patients had biopsy-proved N3 or M1 disease, the stage was recorded, but they were not resected. If there was no evidence of N2 or higher disease, patients underwent thoracotomy, pulmonary resection, and complete thoracic lymphadenectomy. Pathologic review was performed by using standard techniques, and immunohistochemical staining was used in selected cases per the pathologist's discretion.

Definitions

A patient was defined as having unsuspected N2 disease (false-negative result) if the integrated FDG-PET/CT scan did not sug-

gest cancer in any of the N2 nodes but the patient had pathologic proof of metastatic NSCLC cancer in at least one N2 node. Similarly, a patient was defined as having a false-positive result if the integrated FDG-PET-CT scan suggested disease in the N2 nodes but the patient did not have disease in that node on pathologic examination. Similarly, the accuracy of the T status was determined by the false-positive and false-negative results for the T status. For instance, if the FDG-PET suggested a tumor was T2 but after resection it was T3 or T4, this was considered a false-negative result. Thus if the pathologic T status was greater than the PET-predicted T status, it was considered a false-negative result. If

TABLE 3. Efficacy of integrated FDG-PET/CT for the 3 groups

	Never smoked (%) n = 52	Former smokers (%) n = 112	Smokers (%) n = 82	P value*
Overall accuracy	83	80	64	.03
T status (all)	88	84	86	.54
N status (all)	89	72	79	.08
N2	96	75	72	.02
M status	81	85	78	.73

FDG-PET, fluorodeoxyglucose–positron emission tomography. *Compares those who never smoke with current smokers.

Figure 1. Distribution of false-positive and false-negative results between smokers and nonsmokers. Positron emission tomographic/computed tomographic scan results were more often false negative in patients who were current smokers.

the pathologic T status was less than the predicted T status, it was considered a false-positive result. For this calculation, T1 and T2 were considered the same, and T3 and T4 were considered the same. Operative morbidity and mortality is defined as any morbidity or mortality occurring during the hospital stay, within 30 days after discharge, or both.

Statistical Methods

Data were imported from the prospective database (Excel; Microsoft Corp, Seattle, Wash) into an Access database (Microsoft Corp). Analysis was performed with EpiInfo (Centers for Disease Control and Prediction, Atlanta, Ga) and SAS 9.0 (SAS, Cary, NC). Accuracy (defined as the true-negative plus true positive results divided by the sum of all true and false results) was determined for FDG-PET/CT by using the pathology or biopsy results as the gold standard.¹⁰ Categorical values were compared by using analysis of variance, the χ^2 test, or the Fisher exact test. Continuous variables were compared by using the Student *t* test for normally distributed variables and the Wilcoxon test for nonnormally distributed variables.

Results

Patient and Tumor Characteristics

Patient characteristics are shown in Table 1. Significantly more women were nonsmokers ($P = .02$). Nonsmokers had better percentage forced expiratory volume in 1 second ($P = .016$) and percentage diffusing capacity in the lung from carbon monoxide ($P = .037$) values compared with those seen in smokers in this study. Tumor characteristics are

shown in Table 2. Despite the lesions being of similar size, nonsmokers had a significantly greater maximum FDG uptake value (maxSUV) compared with smokers (7.3 vs 4.9, respectively; $P = .026$). The median maxSUV value of N2 lymph nodes (excluding nodes that had a value of <2.5) was higher for nonsmokers compared with that for smokers ($P = .09$). The stage, histology, degree of differentiation, and lymphovascular invasion were similar among smokers and nonsmokers. Operative morbidity and mortality was 19.2% and 2% in those who never smoked, 21.4% and 1% in former smokers, and 29% and 2.4% in current smokers.

Accuracy

Table 3 shows that the accuracy was greater for nonsmokers overall and for T, N, N2, and M comparisons. Differences that achieved statistical significance were those for overall accuracy, almost all nodal status, and N2 lymph node status. Interestingly, even former smokers had less accuracy for the prediction of N2 disease when compared with those who had never smoked ($P = .004$). Figure 1 shows the distribution of the false-positive and false-negative results for smokers, former smokers, and those who had never smoked. Table 4 shows the most common causes of false-positive results. In those who never smoked and former smokers, it was infectious (primarily fungal infections). Smokers had more false-negative results than patients who quit smoking or never smoked. Figure 2 shows the inverse relationship between pack-year smoking history and the accuracy of FDG-PET/CT scanning. Not surprisingly, we noted that as the maxSUV of the primary tumor, lymph node, or both decreased, the overall accuracy for the clinical stage also decreased. Multivariate analysis showed that status of smoking ($P = .026$) and maxSUV ($P = .014$) were independent predictors of FDG-PET/CT accuracy, and pack-years approached statistical significance ($P = .073$).

Discussion

In this prospective study we found that patients with NSCLC who continue to smoke cigarettes at the time of their integrated FDG-PET/CT scan have less accurate clinical staging compared with those who stopped 1 month before and those who never smoked. Additionally, there

TABLE 4. Common causes of false-positive results

	No. (%)	Most common cause of false-positive results (no. of patients)	Second most common cause of false-positive results (no. of patients)
Never smoked	23 (44)	Infection (13)	Granulomatous disease (6)
Quit smoking	40 (36)	Infection (16)	Granulomatous disease (8)
Smokers	23 (28)	Granulomatous disease (12)	Infection (8)

Figure 2. Pack-year smoking history versus accuracy of predicting N2 lymph node disease.

was a dose response associated with the amount of cigarettes smoked and the inaccuracy of integrated FDG-PET/CT scans. As the cigarette pack-year history increased, the accuracy of FDG-PET/CT scans to predict the pathology of N2 lymph nodes decreased ($P = .04$).

We found that smokers have lower maxSUV values of the primary tumor and of the N2 lymph node than those who never smoked. This might be attributed to differences in pathophysiology, which result in higher background FDG tracer values in smokers. A higher background uptake of FDG results in a lower maxSUV value of the pulmonary nodule or the lymph node in question (as shown in Figure 3). The results of several studies, including this one, support the hypothesis that a higher background value might exist in smokers compared with nonsmokers. Miyauchi and Wahl¹¹ showed in 1996 that smokers had significantly lower SUV values of nodules with similar histology when measured at both 50- to 60-minute and 60- to 70-minute intervals compared with their nonsmoking counterparts. Additionally, some studies have postulated explanations for higher background uptake in smokers compared with

nonsmokers. Engel and associates¹² in 1996 found that muscles with greater activity have greater uptake/background compared with muscles at rest. Smokers might use more chest musculature for breathing at rest than nonsmokers given their propensity for chronic pulmonary lung obstruction and emphysema. Additionally, smokers have less blood flow to the lung¹³ and less oxygen exchange, which might cause FDG to accumulate, thus increasing background values. Likewise, smokers in our study had significantly decreased pulmonary function test results and lower overall Eastern Cooperative Oncology Group performance compared with those who never smoked.

The predictive value of integrated FDG-PET/CT is dependent on the maxSUV value of the individual N2 lymph nodes. In our previous study in 2005,¹⁴ we used receiver operating characteristic curves to identify the optimal maxSUV value of mediastinal lymph nodes that best predicted pathology. We found that value to be 5.3 for N2 (mediastinal) lymph nodes. However, in this study we used a lower value of 2.5 so as not to miss N2 disease and thus accepted some false-positive results. Therefore the lower accuracy in smokers for N2 lymph nodes might be due to the increased background activity, which results in a lower maxSUV of the nodes compared with that seen in those who never smoked.

The strengths of this study are the prospective design, the pathology-proved nodal disease with removal instead of biopsy, the fact that the distribution of the histology of the tumors was similar among smokers and nonsmokers, and the fact that clinical staging was performed by one individual for all patients. Limitations of this study are the small number of patients and the fact that there might be undiscovered differences in the molecular characteristics, biochemical characteristics, or both of cancers in smokers compared with nonsmokers. Nonsmokers are more likely to have bronchoalveolar and carcinoid cancers, and because these types of cancers are less likely to metastasize to N2 nodes, this distribution might affect the analysis. The take-

Figure 3. Simplified hypothesis of the affect of smoking on fluorodeoxyglucose-positron emission tomographic uptake. Smokers might have a higher background fluorodeoxyglucose uptake, resulting in a decreased maxSUV of the nodule. In this example, a nodule with a maxSUV of 10 would be reported with a maxSUV of 6 in a smoker and 8 in a nonsmoker because of differences in background uptake. maxSUV, maximum standardized uptake value.

home message from our study is that when integrated FDG-PET/CT is used to clinically stage patients with NSCLC, the clinician should be aware that if the patient is smoking cigarettes at the time of the integrated FDG-PET/CT scan, the maxSUV values might be lower, and the clinical accuracy might be lower as well.

References

1. Cerfolio RJ, Ohja B, Bryant AS, et al. The role of FDG-PET scan in staging patients with nonsmall cell carcinoma. *Ann Thorac Surg.* 2003;76:861-6.
2. Lardinois D, Weder W, Hany TF, et al. Staging of non-small cell lung cancer with integrated positron emission tomography and computed tomography. *N Engl J Med.* 2003;348:2500-7.
3. Aquino SL, Asmuth JC, Alpert NM, et al. Improved radiologic staging of lung cancer with 2-[18]-fluoro-2-deoxy-D-glucose-positron emission tomography and computed tomography registration. *J Comput Assist Tomogr.* 2003;27:479-84.
4. Cerfolio RJ, Bryant AS, Ojha B, et al. Improving the inaccuracies of clinical staging of patients with NSCLC: a prospective trial. *Ann Thorac Surg.* 2005;80:1207-13.
5. Reed CE, Harpole DH, Posther KE, 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;126:1943-51.
6. Nabi HA, Zubeldia JM. Clinical applications of F18-FDG in oncology. *J Nucl Med Technol.* 2002;30:3-9.
7. Mountain CF. Revisions in the International System for Staging Lung Cancer. *Chest.* 1997;111:1710-7.
8. Hawes RH, Gress F, Kesler KA, Cummings OW, Conces DJ Jr. Endoscopic ultrasound versus computed tomography in the evaluation of the mediastinum in patients with non-small-cell lung cancer. *Endoscopy.* 1994;26:784-7.
9. Cerfolio RJ, Bryant AS, Ojha B, et al. Improving the inaccuracies of clinical staging of patients with NSCLC: a prospective trial. *Ann Thorac Surg.* 2005;80:1207-14.
10. Beck J. Likelihood ratios: another enhancement of sensitivity and specificity. *Arch Pathol Lab Med.* 1986;110:685-6.
11. Miyauchi T, Wahl RL. Regional 2-[18F]fluoro-2-deoxy-D-glucose uptake varies in normal lung. *Eur J Nucl Med.* 1996;23:517-23.
12. Engel H, Steinert H, Buck A, et al. Whole-body PET: physiological and artifactual fluorodeoxyglucose accumulations. *J Nucl Med.* 1996;37:441-6.
13. Permutt S, Howell JBL, Proctor DF, Riley RL. Effect of lung inflation on static pressure-volume characteristics of pulmonary vessels. *J Appl Physiol.* 1961;16:64-70.
14. Bryant AB, Cerfolio RJ, Klemm K, et al. The maxSUV of mediastinal lymph nodes on integrated FDG-PET-CT predicts pathology in patients with non-small cell lung cancer. *Ann Thorac Surg.* 2006;82:417-22.

Discussion

Dr David H. Harpole (*Durham, NC*). As usual, very nicely done and nicely presented. I would postulate that the difference in the lymph nodes actually might be due to the histology of the lymph nodes, and it might be worth examining further because smokers have a lot of histiocytes that are just packed with material from the smoke. We have previously investigated this for a correlation with the glut receptors. Therefore it is likely that these nodes are a little hard to read from that standpoint and probably would have a lower maxSUV value just because there are more of these histiocytic nodes, which are full of anthracotic pigment and are not that metabolically active, so that the tumors are a little more dispersed in those nodes than in patients who were never smokers. The other thing I would want to look at in those who never smoked is to see how many germinal centers there are in those nodes and how packed the tumor is and so forth. I think these are excellent data. I am just trying to come up with sort of histopathologic reason why there might be a difference in those nodes, but otherwise this is very nicely done.

Dr Bryant. Thank you, Dr Harpole. We agree with your comments.

Dr Mark J. Krasna (*Baltimore, Md*). I am just wondering about the relationship between the smokers and those who never smoked versus sex. It seemed very apparent on that first slide that you were also looking at a group that had more women. With the new information on women, nonsmokers, and the prevalence of bronchoalveolar carcinoma in that subset, can you perhaps give us some explanation on either how this might explain the findings or how you might go about trying to get rid of that possible bias in a follow-up study?

Dr Bryant. Thank you, Dr Krasna, for your question. There is a propensity for bronchoalveolar carcinoma in the population who never smoked in this series. In our study we used a multivariate analysis model in which we controlled for sex, and in doing so, we still found that the status of smoking, as well as the maxSUV value, were independent predictors of accuracy on PET scanning. Dose response was very close to being an independent predictor, with a *P* value of .07, which is just higher than the .05 level, the usual arbitrary cutoff to indicate statistical significance.