Advances in positron emission tomography technology have increased the need for surgical staging in non–small cell lung cancer

Benjamin Enoch Lee, MD, a Derek von Haag, MD, a Teri Lown, RN, OCN, b Derick Lau, MD, PhD, c Royce Calhoun, MD, a and David Follette, MD a

Objectives: Pretreatment staging of patients with non–small cell lung cancer is critically important in determining an appropriate treatment plan. As positron emission tomography (PET) and computed tomography (CT) are proven complementary modalities in clinical staging, recent advances in PET technology have brought forth integrated PET/CT as the new standard. We tested the hypothesis that improvements in PET technology have not increased the sensitivity or specificity of PET in the staging of non–small cell lung cancer to an extent that surgical staging is no longer required.

Methods: This is a retrospective, single-institution review of 336 patients from 1995 to 2005 with biopsy-proven non–small cell lung cancer who underwent [18F] fluoro-2-deoxy-D-glucose–PET before mediastinal lymph node sampling by cervical mediastinoscopy or thoracotomy. Clinical records, histopathologic reports, and PET findings were reviewed. Data were analyzed by the Pearson χ² test.

Results: Within the study population, 210 patients had routine PET and 126 had integrated PET/CT. For detecting mediastinal metastases the sensitivities of PET versus integrated PET/CT were 61.1% versus 85.7% (P < .05), specificities were 94.3% versus 80.6% (P < .001), positive predictive values were 68.8% versus 55.8%, negative predictive values were 92.1% versus 95.2%, and overall accuracy was 88.6% versus 81.7%.

Conclusions: Improvements in PET technology have increased integrated PET/CT sensitivity at the cost of significantly decreased specificity. Although it may appear that integrated PET/CT incurs fewer false negative results, the dramatic increase in false positive results reinforces the notion that integrated PET/CT should be used only as an adjunct to clinical staging and that surgical staging remains the gold standard in non–small cell lung cancer.

Non–small cell lung cancer (NSCLC) remains the leading cause of cancer death in the United States today.1 Given the recent advances in therapeutic options and the changing treatment algorithms, accurate pretreatment staging in NSCLC is of paramount importance in formulating an appropriate treatment plan. Patients without mediastinal lymph node disease (N0 or N1) have a better prognosis and have traditionally been treated with surgical resection alone.2 Recent evidence has demonstrated a modest improvement in survival with the addition of adjuvant chemotherapy for patients with stage II or IIIA disease.3,4 On the other hand, patients with gross mediastinal disease (N2 or N3) are commonly treated with definitive chemoradiotherapy.5,6

Whole-body positron emission tomography (PET) with [18F] fluoro-2-deoxy-D-glucose (FDG) has rapidly become accepted as the standard noninvasive modality for staging in patients with NSCLC.7–11 Although PET has been shown to be...
superior to computed tomography (CT) for staging of NSCLC, in reality PET and CT are complementary modalities whose combined diagnostic value is superior to either study alone. Accordingly, technologic advances have introduced integrated PET/CT as the newest modality in the armamentarium of cancer staging.

Given the novelty of PET/CT, the number of studies comparing PET and PET/CT in NSCLC are limited, but growing. Recent studies have documented superior accuracy with integrated PET/CT over PET alone in overall staging and diagnosis of NSCLC. However, these studies have also introduced controversy as to whether PET/CT is superior to PET for nodal staging of the mediastinum. Because of a greater than 95% negative predictive value, current practice accepts a negative PET result without the need for surgical confirmation. By contrast, a positive PET scan requires surgical confirmation because of the high false positive rate from coexistent inflammatory or infectious processes. Practically, however, it is not infrequent that patients are treated with neoadjuvant therapy for suspected N2 disease or definitive chemoradiotherapy for N3 disease solely on the basis of a positive PET result. Previous studies showed that surgical staging was still required because of a relatively high false positive rate with PET in mediastinal staging. We contend that integrated PET/CT also has not improved specificity to replace surgical staging as the sole diagnostic tool in NSCLC. Accordingly, the purpose of this study was to review our experience and compare the diagnostic accuracy between PET and integrated PET/CT for nodal staging in NSCLC.

**Patients and Methods**

**Patient Selection**

A retrospective review was performed on all patients who underwent surgical mediastinal lymph node biopsy by cervical mediastinoscopy, anterior mediastinotomy, thoracotomy or a combination of these methods, between January 1995 and December 2005 on the Thoracic Surgery Service at the University of California, Davis Cancer Center. Only newly diagnosed patients with biopsy-proven NSCLC and preoperative staging PET scans were included. Patients were segregated into two study groups: standard PET versus simultaneously acquired integrated PET/CT. The PET scans were then compared with the reference standard of pathologic results to determine sensitivity, specificity, positive and negative predictive values, and accuracy. This study was approved by the Institutional Review Board at the University of California at Davis Medical Center.

**PET and Integrated PET/CT Imaging**

All PET studies were performed after patient fasting for a minimum of 4 hours. PET images were obtained with a dedicated PET system (ECAT EXACT 921; CTI, Knoxville, Tenn). PET/CT images were obtained with an integrated PET/CT scanner (Discovery LS; GE Medical Systems, Waukesha, Wis; or ECAT Revive XVI; CTI, Knoxville, Tenn). Whole-body scans were obtained 30 to 60 minutes after intravenous injection of 10 to 20 mCi of FDG. For PET imaging, projection and tomographic images in the axial, coronal, and sagittal planes were reconstructed both with and without attenuation correction. For PET/CT imaging, simultaneously acquired CT data were used for attenuation correction. All studies were read by dedicated nuclear medicine physicians with a specialty in interpreting PET scan images. Clinical histories and pertinent CT scans were available for review. Intraobserver variability was not assessed. Mediastinal lymph nodes were read as positive if their activity was definitely above the surrounding mediastinal activity and not according to standard uptake values (SUV). Only patients with ipsilateral (N2) or contralateral (N3) mediastinal disease were considered to have positive results for this study.

**Mediastinal Lymph Node Staging**

Extended mediastinal lymph node staging was completed in all patients by cervical mediastinoscopy, anterior mediastinotomy, or thoracotomy. In patients with normal mediastinoscopy results, thoracotomy followed typically within 14 days. The results of PET and CT scanning were available to the surgeon at the time of resection. All visible and technically feasible lymph nodes were removed and were annotated according to the revised International Staging System. Pathologic reports were reviewed to determine whether any mediastinal lymph nodes contained cancer. Only patients with pathologic disease in lymph nodes that would have been accessible by mediastinoscopy (stations 2, 3, 4, and 7), mediastinotomy (stations 5 and 6), right thoracotomy (stations 2, 4, 7, and 9), or left thoracotomy (stations 4, 5, 6, and 9) were considered positive in this study. There were no changes in surgical routine during the study period.

**Statistics**

Pathologic findings served as the “gold standard.” Sensitivity, specificity, positive predictive value, negative predictive value, and accuracy were calculated at the patient level. Exact binomial confidence intervals were determined for each. Comparisons between PET and PET/CT diagnostic test characteristics were made with the $\chi^2$ test for independent groups. Differences between the groups on demographic and baseline characteristics were assessed by the $\chi^2$ or $t$ test. Statistical analysis was carried out with SAS version 9.1 for Windows (SAS Institute, Inc, Cary, NC).
Results

Between January 1995 and December 2005 at the University of California, Davis Cancer Center, 629 patients underwent surgical mediastinal lymph node biopsy with an average of 4 lymph node stations sampled. Of the 293 patients excluded from this study, 269 did not have pathologic diagnosis of NSCLC, 16 patients did not have a preprocedure PET scan, 5 patients had previous staging procedures, and 3 patients did not have lymph node tissue on biopsy.

The remaining 336 patients were segregated into standard PET (n = 210) and PET/CT (n = 126). Table 1 shows the population demographics of the two patient groups. The standard PET group had 44% men and 56% women with a mean age of 65.1 years (range 32-86 years). The PET/CT group had 42% men and 58% women with a mean age of 67 years (range 37-86 years). The accuracy of mediastinoscopy in the PET group compared with the PET/CT group was 82.6% (50/61) and 97.7% (43/44), respectively (data not shown).

The prevalence of mediastinal metastases by surgical staging was 17.1% (36/210) in the PET group compared with 22.2% (28/126) in the PET/CT group (P = not significant [NS]). For detecting mediastinal metastases the sensitivities of standard PET versus PET/CT were 61.1% versus 85.7% (P < .05); specificities were 94.3% versus 80.6% (P < .001); positive predictive values were 68.8% versus 55.8% (P = NS); negative predictive values were 92.1% versus 95.2% (P = NS); and overall accuracy was 88.6% versus 81.7% (P = NS) (Tables 2 to 4).

There were 4.8% (10/310, 95% confidence interval [CI] 2.3%-8.6%) false positive results in the PET group compared with 15.1% (19/126, 95% CI 9.3%-22.5%) in the PET/CT group. Within the PET group, 8 of 10 patients had coexistent inflammatory or infectious etiologies. One patient may have had a sampling error inasmuch as the mediastinoscopy results were negative, but the patient was found to have stage IV disease at thoracotomy. One patient was believed to have an incorrect interpretation of primary tumor activity adjacent to the mediastinum. Within the PET/CT group, 17 of 19 patients had coexistent inflammatory, neoplastic, or infectious etiologies. Two patients had no identifiable coexistent pathologic processes.

Fifty-one (24%) of 210 patients in the PET group and 44 (35%) of 126 patients in the PET/CT group underwent a mediastinoscopy followed by thoracotomy owing to node-negative disease (Table 5). The indications for mediastinoscopy in the PET versus the integrated PET/CT subgroups were a false positive PET scan (10/51, 20%, vs 19/44, 43%), a false positive CT (27/51, 53%, vs 22/44, 50%), or suggestion of either N1 disease or tumor proximity to the mediastinum (22/51, 43%, vs 17/44, 39%), respectively (P < .01) (Table 6).

PET/CT had a false negative rate of 4.8% (4/83) when compared with PET (7.9%, 14/178). The 4 patients with false negative results in the PET/CT group had metastatic disease on standard pathologic evaluation and thus did not have micrometastatic disease on histopathologic examination nor did they lack significant uptake in their primary tumors or mistaken evidence of N1 disease on PET/CT. Three of the primary tumors were clinically and pathologically T1 and the other was T2. The average length of time from PET/CT to resection was 33.8 days.

Discussion

Our results indicate that integrated PET/CT has not improved the overall accuracy of mediastinal staging in NSCLC.
NSCLC over standard PET imaging. Improved sensitivity in the detection of N2/N3 disease with PET/CT has come at the cost of significantly worsened specificity. The drastic increase in false positive results reinforces the continued need for surgical staging in the treatment of NSCLC.

FDG-PET imaging is the most accurate noninvasive staging modality for NSCLC available today.9,21 By using the higher rate of glycolysis in malignant cells compared with normal surrounding tissues, FDG-PET allows a physiologic assessment of tumor activity.22 Despite the high incidence of false positive results, meta-analysis of PET utility in mediastinal staging has produced a pooled sensitivity of 84% and a specificity of 89%, which compares favorably with CT (sensitivity of 57% and specificity of 89%), which compares favorably with CT (sensitivity of 57% and specificity of 82%).9 Although superior, the inherent limitations in anatomic precision of PET imaging have led to the evolution of combining PET imaging with CT scanning. Initial efforts using computer software to create fusion images were met with alignment difficulties from using images taken at different time points.23,24 The advent of integrated PET/CT has made simultaneous image acquisition possible, thus ameliorating the problem.

Many studies have indicated the overall superiority of integrated PET/CT over CT alone,25 PET alone,13-15 and visually corrected or fused PET/CT15,26,27 in NSCLC. However, with regard to nodal staging of the mediastinum in NSCLC, there does not appear to be a consensus agreement in the utility of integrated PET/CT. A review of the recent literature reveals that most studies comparing PET/CT with PET do not reach statistical significance with respect to sensitivity or specificity because of small sample size. In one of the larger series, Cerfolio and associates14 compared PET/CT with PET and showed an overall increased accuracy in both T and N staging. However, in the identification of N2 disease, increased accuracy (96% vs 93%) was a result of improved sensitivity (69% vs 62%) and came at the cost of worsened specificity (94% vs 97%), which is partly supported by our results.

Logically, this makes sense. Increased sensitivity is the result of identifying subtle lesions smaller than 2 cm, which may have been lost in the background with standard PET imaging. These lesions are now visually correlated and more easily identified with the aid of CT imaging. This, however, has not resolved the established difficulty of PET with false positive results from inflammatory or infectious diseases.28,29

What is even more interesting was the significant impact the increased false positive results had on the types of operations performed. The PET and PET/CT groups had significantly different proportions of operations performed, which were primarily manifested by a higher rate of thoracotomy in the PET group compared with a higher rate of mediastinoscopy followed by thoracotomy in the PET/CT group. This disparity appears to be due to a significantly increased number of patients with false positive PET/CT scans, who, after a negative mediastinoscopy are able to proceed to thoracotomy. In contrast, neither a false positive CT scan, presence of N1 disease, nor tumor proximity to the mediastinum resulted in any significant difference. This striking observation highlights the increased need for surgical confirmation of a positive integrated PET/CT scan as this will allow a significant number of patients to proceed to curative resection rather than being subjected to neoadjuvant or definitive chemoradiation therapy.

| TABLE 4. Efficacy of mediastinal staging by PET versus PET/CT |
|---------------------------------|-----------------|-----------------|-----------------|
|                                 | Standard PET (n = 210) | Integrated PET/CT (n = 126) | P value* |
| Sensitivity (%), 95% CI         | 61.1 (43.5-76.9) | 85.7 (67.3-96.0) | .0299 |
| Specificity (%), 95% CI         | 94.3 (89.7-97.2) | 80.6 (71.4-87.9) | .0005 |
| Positive predictive value (%), 95% CI | 68.8 (50.0-83.9) | 55.8 (39.9-70.9) | .2552 |
| Negative predictive value (%), 95% CI | 92.1 (87.2-95.6) | 95.2 (88.1-98.7) | .3658 |
| Accuracy (%), 95% CI            | 88.8 (87.2-95.6) | 81.7 (88.1-98.7) | .0808 |

PET, Positron emission tomography; CT, computed tomography. CI, confidence interval. *P value for χ² test.

| TABLE 5. Analysis of operations performed |
|------------------------------------------|-----------------|-----------------|
| Characteristic                           | Standard PET (n = 210) | Integrated PET/CT (n = 126) |
| Mediastinoscopy or mediastinotomy (No., %) | 33 (16) | 25 (20) |
| Mediastinoscopy followed by thoracotomy (No., %) | 51 (24) | 44 (35) |
| Thoracotomy (No., %)                      | 126 (60) | 57 (45) |

PET, Positron emission tomography; CT, computed tomography. P value for χ² test = .029.

| TABLE 6. Analysis of patients undergoing mediastinoscopy followed by thoracotomy |
|------------------------------------------|-----------------|-----------------|
| Characteristic                           | Standard PET (n = 51) | Integrated PET/CT (n = 44) | P value* |
| False positive PET or PET/CT (No., %)    | 10 (20) | 19 (43) | .003 |
| False positive CT (No., %)               | 27 (53) | 22 (50) | .775 |
| Proximity or N1 disease (No., %)         | 22 (43) | 17 (39) | .796 |

PET, Positron emission tomography; CT, computed tomography. A significantly increased number of patients were able to proceed to thoracotomy after a node negative mediastinoscopy owing to a false positive integrated PET/CT scan. Also, the percentages do not total 100% as some patients had multiple indications for mediastinoscopy. *P value for χ² test.
This study also affirms that integrated PET/CT maintains a low false negative rate and is sufficient evidence to rule out mediastinal spread. Recent evidence has suggested that T2 tumors are more likely to harbor occult N2 disease in patients with clinical stage I disease. However, owing to a limited sample size, this study was not able to affirm or deny those findings. Thus, although the spatial resolution with integrated PET/CT is much improved from standard PET, detection of subcentimeter lymph nodes may still pose an obstacle in ruling out metastatic disease.

Finally, the future of pretreatment staging in NSCLC is rapidly changing. As recent trials using adjuvant platinum-based chemotherapy in patients with early-stage NSCLC have shown significantly improved survival, trials are now investigating the use of similar therapy in the neoadjuvant setting. The real impact of this will be felt in the increased need to more accurately stage the disease before therapy. Although not addressed in this study, PET/CT has been shown to be more accurate in segmenting patients with N0 versus N1 disease. However, the high false positive rate in PET/CT will still mandate confirmation of N1 disease. Emerging invasive technologies, such as the use of fine needle aspiration with endoscopic ultrasound and endobronchial ultrasound are still being evaluated and may have significant benefit in confirming N1 disease. Suggestion has even been made that these technologies will supplant the utility of mediastinoscopy in the future. More likely, however, will be the future development of modality staging algorithms to address the changing treatment algorithms in NSCLC. Until then, the current evidence still supports the continued use of mediastinoscopy as the gold standard for confirmation of mediastinal disease in NSCLC.

References

Discussion

Dr John D. Mitchell (Denver, Colo). This study details your institution’s experience with PET and subsequently PET/CT for staging the mediastinum in the setting of NSCLC. Using integrated PET/CT, you found improved sensitivity but reduced specificity in detecting N2 disease. The negative predictive value remained high, greater than 95%, and the overall accuracy of the imaging test declined somewhat with PET/CT to 82%. On the basis of the increased false positive rate you saw with PET/CT, you advocate for the continued need for mediastinoscopy to best stage the mediastinal extent of disease. I agree with the ongoing need for surgical staging in the mediastinum, and I have a few questions for you.

First, we have had some issues at our institution with our nuclear medicine physicians interpreting the CT part of the PET/CT. In your manuscript you report that dedicated nuclear medicine colleagues read these studies. Do you believe that they have adequate body imaging experience or have you seen similar problems at your institution? Could this account for some of the changes you saw between PET and PET/CT?

Dr Lee. I agree and this is actually a good point that was brought up during our morning session. We have two centers that do the vast majority, greater than 95%, of our integrated PET/CT imaging at our institution, and each center essentially has dedicated physicians specialized in nuclear medicine reading who are also specialized in body chest CT reading. At our institution we do not have that detriment of a lack of expertise in CT imaging. However, since the CT images acquired during PET/CT are without contrast, any radiologist will have a more difficult time comparing those images versus a contrasted CT scan. However, most of these patients also have a previous contrasted CT scan, which also is used during the time of the reading.

Dr Mitchell. Second, what constitutes a positive study at your institution? In your manuscript you described a study as positive if the degree of activity was “definitely above the surrounding mediastinum.” Do you use a specific SUV as a cutoff for a positive study? It is going to have a tremendous effect on the sensitivity and specificity of the test.

Dr Lee. That is a good question that was actually discussed this morning. At our institution, our radiologists typically will assign an SUV to our primary tumors but have not yet accepted the practice of assigning SUVs to lymph nodes. It is our common practice to use background uptake as a baseline level and designate as “positive” anything with increased SUV compared with background. Because of this we probably were able to have an increased sensitivity. Likewise, because of this, the specificity is much lower, which leads to increased false positive results.

Dr Mitchell. On the basis of your results, have you thought about reinterpreting your data that would adjust your sensitivity and specificity based on the SUV value?

Dr Lee. That is something that we will probably want to go back and look at. Then if we can stratify on the basis of SUV we might be able to see exactly whether there is a cutoff. Dr Cerfolio has identified around 5.2 as being that magic number at his institution, and I know there is some variance from institution to institution as well. I think that is something that we definitely will look forward to looking into.

Dr Mitchell. Next, the negative predictive value in your study remains high, greater than 95%, with integrated PET/CT. On the basis of your results, what do you advocate in terms of surgical staging in the mediastinum if the PET/CT results are negative?

Dr Lee. If the patient has a peripheral lesion that patients should be assumed to have, and if the PET/CT results are negative, patients should be assumed to have a normal mediastinum and should proceed to direct thoracotomy and resection. Patients with central tumors, evidence of N1 disease, should be more closely evaluated with mediastinoscopy first. If the patients have large bulky central lesions that may require pneumonectomy, it is to their benefit to clearly identify or rule them out from having mediastinal disease before pneumonectomy.

Dr Mitchell. Finally, others have reported increased accuracy in staging the mediastinum, as you alluded to, using an integrated PET/CT, but your data suggest otherwise. How do you account for the differences in the studies?

Dr Lee. We also touched on this a little bit. It has to do with how you assign what is a positive integrated PET/CT. If the SUV or if your way of identifying or calling a positive study is set too low, you will have a higher number of false positive results. There has also been evidence that geographic differences exist. In our part of the country, we actually have a much higher incidence of both sarcoid disease and histoplasmosis. That may play into the fact that we had a much higher incidence of false positive results.

Dr Robert Cerfolio (Birmingham, Ala). First of all, congratulations on your fine work. I encourage you to do these types of studies and I invite you to add the max-SUV and cutoff values because they will completely change your data and your results—whether the node is positive or not. I think you need to set the bar for the max-SUV of the lymph node and objectively call it positive or negative and then prove whether you are right or wrong. Although your study and our study have somewhat different conclusions, I do not think they are all that different. What I was hoping to see was a specific T and an N and a specific nodal station analysis, as we have described in several papers. Was that included in your manuscript but left out of your presentation for lack of time, or was that not done?

Dr Lee. That actually was not done. The reason is that our nuclear medicine readouts are not assigned to say we have an R4 node that is positive. They will typically just say the mediastinum is positive.

Dr Cerfolio. I think if you are going to look at the accuracy of
the TNM, you have to look at the accuracy for each nodal station. Ask the nuclear radiologists what the T is and what the N is and which N is positive to see if they are right.

Dr Lee. I totally agree. It may be that we will have to change our practice at our institution to start doing that.

DOCTOR. I would encourage you to take this to the pulmonology meetings, to the medical oncology and radiation oncology meetings, because they are reading those tests. They are bypassing you and are treating these patients as if they have advanced disease. The specificity question is vital, and I would strongly encourage you to take this information to these people. Tissue is the issue, and we have to tell them that.