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

Carolyn E. Reed, MD, a David H. Harpole, MD, b Katherine E. Posther, MD, c Sandra L. Woolson, MPH, d Robert J. Downey, MD, d Bryan F. Meyers, MD, e Robert T. Heelan, MD, f Homer A. Macapinlac, MD, f Sin-Ho Jung, PhD, i Gerard A. Silvestri, MD, j Barry A. Siegel, MD, k and Valerie W. Rusch, MD l

Objectives: The American College of Surgeons Oncology Group undertook a trial to ascertain whether positron emission tomography with 18F-fluorodeoxyglucose could detect lesions that would preclude pulmonary resection in a group of patients with documented or suspected non–small cell lung cancer found to be surgical candidates by routine staging procedures.

Methods: A total of 303 eligible patients registered from 22 institutions underwent positron emission tomography after routine staging (computed tomography of chest and upper abdomen, bone scintigraphy, and brain imaging) had deemed their tumors resectable. Positive findings required confirmatory procedures.

Results: Positron emission tomography was significantly better than computed tomography for the detection of N1 and N2/N3 disease (42% vs 13%, P = .0177, and 58% vs 32%, P = .0041, respectively). The negative predictive value of positron emission tomography for mediastinal node disease was 87%. Unsuspected metastatic disease or second primary malignancy was identified in 18 of 287 patients (6.3%). Distant metastatic disease indicated in 19 of 287 patients (6.6%) was subsequently shown to be benign. By correctly identifying advanced disease (stages IIIA, IIIB, and IV) or benign lesions, positron emission tomography potentially avoided unnecessary thoracotomy in 1 of 5 patients.

Conclusions: In patients with suspected or proven non–small cell lung cancer considered resectable by standard staging procedures, positron emission tomography can prevent nontherapeutic thoracotomy in a significant number of cases. Use of positron emission tomography for mediastinal staging should not be relied on as a sole staging modality, and positive findings should be confirmed by mediastinoscopy. Metastatic disease, especially a single site, identified by positron emission tomography requires further confirmatory evaluation.

Accurate staging of non–small cell lung cancer (NSCLC) is essential to the ability to offer a patient the most effective available treatment and the best estimate of prognosis. The most significant recent advance in the staging of NSCLC has been the use of positron emission tomography (PET) with the glucose analog 18F-fluoro-2-deoxy-D-glucose (FDG). FDG-PET provides information regarding the functional activity of a malignant lesion rather than strictly anatomic information as is provided by computed tomography (CT).

The use of PET for the evaluation of solitary pulmonary nodules is now well established. Gould and colleagues performed a meta-analysis on 40 studies inclusive of 1474 patients who had solitary lung nodules evaluated by PET. The overall sensitivity for detection of malignancy was 96.8%, and the specificity was 77.8%.
Evaluation of mediastinal nodal disease is crucial for staging NSCLC. The accuracy of CT for predicting malignant mediastinal lymphadenopathy is dismally low (sensitivity 57% and specificity 82%), as reported in a meta-analysis of 18 studies and 1045 patients by Toloza and coworkers. This meta-analysis documented that PET has improved sensitivity and specificity for the detection of malignancy in the mediastinum (84% and 89%, respectively). These results prompted the American College of Chest Physicians in its recent lung cancer guidelines to recommend FDG-PET for mediastinal staging.

The use of PET for the detection of distant metastases has been insufficiently studied, although reports have documented unexpected metastatic disease in 10% to 20% of patients who are otherwise deemed candidates for surgery. In addition, PET does not accurately detect metastases to the brain because of the high uptake of FDG by normal gray matter.

To validate the reported findings in single-institution studies and to clarify the role of FDG-PET in the staging of patients with potentially resectable NSCLC (stages I, II, and selected IIIA), the American College of Surgeons Oncology Group (ACOSOG) undertook the multi-institutional trial Z0050. The primary objective was to ascertain whether FDG-PET could detect lesions that would preclude pulmonary resection in patients found to be surgical candidates after standard imaging procedures.

Material and Methods

Regulatory Requirements

Investigators registering patients in the Z0050 trial were required to be members in good standing of ACOSOG, and each clinical site had active status with ACOSOG. The Z0050 protocol underwent institutional review board approval at all clinical sites, and subjects gave informed consent before participation in the study. Before enrolling patients in the study, each participating institution was required to submit its protocol for PET imaging along with images from 3 consecutive PET studies for evaluation and approval by the ACOSOG PET Quality Assurance Committee.

Inclusion Criteria

Patients who had (1) a histologically or cytologically confirmed, newly diagnosed, untreated, single-lesion NSCLC; (2) a single lung lesion and a biopsy-proved unilateral mediastinal lymph node metastasis, or (3) a clinical presentation strongly consistent with primary bronchogenic carcinoma; who were medically fit; and who were deemed surgical candidates (stage I, II, or IIIA disease) after standard imaging procedures were eligible. The eligibility criteria were designed to mimic standard clinical practice. Staging procedures had to be performed within 60 days before registration. Standard staging included CT of the chest and upper abdomen (including the adrenals), bone scintigraphy, and contrast-enhanced CT or magnetic resonance imaging (MRI) of the brain. Patients with previous cancers were eligible if curative therapy had been undertaken, there was no evidence of malignancy in the past 5 years, and there was low risk for recurrence. Patients were excluded if they had prior PET for evaluation of NSCLC, evidence by standard staging of stage IIIB or IV disease, or uncontrolled diabetes mellitus.

Clinical and Radiologic Assessments

Clinical assessment included an extended history and physical examination. CT of the chest and upper abdomen (including the adrenals) was performed with fourth-generation scanners capable of high resolution, wide-dynamic range, and rapid screening. Intravenous contrast agent was administered (unless contraindicated or if the lesion was peripheral and there was no evidence of adenopathy). Whole-body bone scintigraphy was performed by standard methods. Scintigraphic abnormalities were further evaluated by appropriate radiologic examinations (radiographs, CT, or MRI) or biopsy or both to exclude metastatic disease. A scintigraphic abnormality with a normal radiograph was considered metastatic disease unless biopsied or otherwise explained by MRI or clinical correlation. CT or MRI of the brain was performed before and after contrast agent administration (unless use of contrast material was contraindicated).

Study Schema

Patients who were clinically staged with I, II, or IIIA disease after standard staging procedures and who were surgical candidates were then registered in the Z0050 trial. The study schema is shown in Figure 1.

FDG-PET

Full-ring dedicated scanners with bismuth germanate or sodium iodide (NaI) detectors and with manufacturer-quoted in-plane spatial resolution of less than 6 mm were used. Integrated PET and CT scanners were not used in this trial. The scanners underwent quality control evaluation on each day imaging was performed. Patients fasted for at least 4 hours before the PET study, but they were well hydrated. The blood glucose concentration was determined immediately before FDG administration, and the study was not performed if the blood glucose concentration was greater than 200 mg/dL. Intravenous fluids, intravenous furosemide, and a Foley catheter to minimize urinary tract FDG activity were recommended, but optional, if filtered backprojection was used for image reconstruction. Patients in whom a Foley catheter was not placed voided before imaging. FDG was obtained from a commercial supplier or was prepared and quality controlled by the routine method in use at the individual sites in a manner consistent with state and federal regulations. The radiochemical purity of the FDG was required to be greater than 90%. With a dedicated bismuth germanate PET system, the FDG dose was 0.14 to 0.21 mCi/kg, with a minimum dose of 10 mCi. With an NaI PET system, a dose of 0.07 mCi/kg was used. FDG was administered intravenously with the patient in a quiet, darkened room, where the patient remained resting comfortably until the time of imaging.

PET imaging of the body was performed first, beginning approximately 45 to 60 minutes after FDG injection. The region imaged extended from the upper/mid-neck to the upper thigh. A series of 5 or 6 overlapping transmission scans (of 2-3 minutes duration) were obtained with a rotating 68Ge/68Ga rod source for each emission scan. A segmentation algorithm was used for atten-
uation correction. With dedicated NaI PET systems, attenuation correction was performed with a 137Cs source in accordance with the manufacturer’s recommendations. Then, 5 or 6 corresponding emission images (each 5-15 minutes in duration depending on the patient’s size) were performed. The PET images were reconstructed both with and without attenuation correction by standard vendor-provided reconstruction algorithms, using either filtered back projection with a Hann filter (frequency cutoff 0.6 × Nyquist = 0.3 cycles/pixel) or ordered-subset expectation maximization with an appropriate filter. After body imaging, a 3-dimensional (10-minute duration) or 2-dimensional (30-minute duration) emission PET study of the brain was performed. A calculated attenuation correction was applied to brain-emission data. The PET images of the brain were reconstructed by filtered back projection with use of a Hann filter (frequency cutoff 0.4 × Nyquist = 0.2 cycles/pixel).

An experienced nuclear medicine physician from each participating site interpreted the PET images. The emission scans were reviewed in axial, coronal, and sagittal planes and as maximum pixel intensity reprojection images, along with transmission images to facilitate anatomic localization of abnormal activity. The images were initially interpreted without knowledge of the results of previously obtained CT, other imaging studies, or surgical staging procedures. The images were then reinterpreted with the help of the CT and the other available conventional imaging studies. The interpreter’s degree of suspicion for cancer at the primary site, in locoregional nodes, and at distant metastatic sites was recorded on a standard form with use of a 5-point ordinal categoric scale with the following categories: 4 = definitely abnormal, 3 = probably abnormal, 2 = indeterminate, 1 = probably normal, and 0 = definitely normal. Locoregional nodes were evaluated in relation to mediastinal blood-pool activity (grade 3 and 4 uptakes were slightly and substantially greater than blood-pool activity, respectively). A score of 3 or 4 was considered positive for final data analysis.

The protocol required that mediastinal abnormalities detected by PET be histologically confirmed with mediastinoscopy, mediastinotomy, thoracoscopy, endoscopic ultrasonography with fine-needle aspiration, or thoracotomy. PET-detected hepatic lesions required biopsy or fine-needle aspiration for cytology to confirm metastatic disease. Benign-appearing cysts or hemangiomas could be confirmed by either MRI or ultrasonography. PET-positive adrenal lesions required biopsy. Osseous abnormalities seen on PET had to be evaluated by appropriate imaging studies (plain

---

**Figure 1. Schema of the American College of Surgeons Oncology Group Z0050 Trial.**

---

The Journal of Thoracic and Cardiovascular Surgery • Volume 126, Number 6 1945
TABLE 1. Final histology of eligible patients

<table>
<thead>
<tr>
<th>Histology</th>
<th>No. patients</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adenocarcinoma</td>
<td>122</td>
<td>40.3</td>
</tr>
<tr>
<td>Squamous cell carcinoma</td>
<td>78</td>
<td>25.7</td>
</tr>
<tr>
<td>Large cell carcinoma</td>
<td>12</td>
<td>4.0</td>
</tr>
<tr>
<td>Non–small cell carcinoma, undesignated</td>
<td>13</td>
<td>4.3</td>
</tr>
<tr>
<td>Bronchoalveolar carcinoma</td>
<td>11</td>
<td>3.6</td>
</tr>
<tr>
<td>Metastatic (non-lung)</td>
<td>7</td>
<td>2.3</td>
</tr>
<tr>
<td>Mixed adenocarcinoma/bronchoalveolar carcinoma</td>
<td>3</td>
<td>1.0</td>
</tr>
<tr>
<td>Adenosquamous carcinoma</td>
<td>3</td>
<td>1.0</td>
</tr>
<tr>
<td>Small cell carcinoma</td>
<td>2</td>
<td>0.7</td>
</tr>
<tr>
<td>Other</td>
<td>17</td>
<td>5.6</td>
</tr>
<tr>
<td>Benign disease</td>
<td>35</td>
<td>11.6</td>
</tr>
</tbody>
</table>

radiographs, CT, MRI, or repeat bone scintigraphy) or biopsy or both.
A lesion reported as a probable metastasis by PET or standard imaging studies but not biopsied was considered a false-positive result if the lesion(s) remained unchanged at 6-month follow-up. If surgical resection was performed, all accessible bronchopulmonary, hilar, and mediastinal lymph nodes were removed and labeled according to the American Thoracic Society lymph node mapping system. Mediastinal lymph nodes for right-sided lesions included 2R, 4R, 7 (8 and 9 if present), and 10R. For left-sided lesions, nodal stations included 5, 6, 7 (8 and 9 if present), 10L, and 4L, if accessible.
The conventional and PET imaging studies from the first 3 patients at each site and a 20% random sample of subsequent patients were reviewed by 3 radiology and nuclear medicine investigators to document that these studies were performed according to protocol specifications.

**Follow-Up**
Follow-up was not required if pulmonary resection was not performed. If surgical resection occurred, follow-up of patient and disease status was required at 6 months.

**Statistical Analysis**
The primary objective of the ACOSOG Z0050 trial was to ascertain whether FDG-PET scanning could detect lesions that would preclude pulmonary resection in patients deemed to be surgical candidates by standard imaging procedures. The study was designed to assess the utility of FDG-PET using data from a clinical construct in which the standard staging procedures had to be performed and interpreted before the use of PET. The design for the primary objective was such that the main unit of measurement was the proportion of patients with negative findings ascertained by standard staging procedures who were subsequently found to have surgical contraindication as a result of the FDG-PET findings. This proportion is therefore a simple measure of the utility of FDG-PET. A true proportion less than 5% was considered clinically insignificant, and a true proportion more than 10% was considered clinically significant. The study was designed to test the null hypothesis that the data are consistent with FDG-PET contraindicating surgery in less than 5% of the patients against the alternative hypothesis that the data are consistent with FDG-PET contraindicating surgery in 10% or more of the patients. A 2-stage accrual design was used to follow for early termination if there was strong early evidence to support the null hypothesis.6

In this 2-stage design, if 5 or fewer patients of the first 120 accrued to the study had PET findings that contraindicated surgery, the study would have been terminated. Otherwise, the study would continue to a second stage to accrue a total of 235 patients. This design has a type I error of 5% and a power of 90%.

**Results**
From January 2000 to December 2002, 445 patients were registered in the Z0050 trial from 22 institutions (Appendix 1). Seventy patients were from a single institution using a consent form that did not follow ACOSOG-mandated institutional review board guidelines. These patients were eliminated from the analysis when this regulatory problem was discovered by an ACOSOG audit of the site. Of the remaining 375 registered patients, 303 were deemed eligible, 12 were formally withdrawn by their site, and 60 were declared ineligible at the time of internal ACOSOG review (16%). Ineligible patients included those with nonresectable disease documented by routine staging (14 patients), missing or outdated staging procedures (15 patients), no adrenal assessment on CT (9 patients), glucose level undocumented or greater than 200 mg/dL (9 patients), regulatory violations (6 patients), and a variety of other reasons (7 patients). Of the 303 eligible patients, 302 were evaluable for PET staging of T and N status (1 patient underwent esophageal biopsy after PET and did not undergo operative staging as the result of a diagnosis of invasive esophageal cancer). PET did not include a separate set of brain images in an additional 15 patients as specified in the protocol, leaving 287 patients who were ultimately evaluable for metastatic disease.
The mean age of eligible patients was 68.2 years (range 43.2–93.4 years); 163 were men and 140 were women. The final histology is shown in Table 1. Of the benign lesions in 35 patients, 26 were confirmed pathologically and 9 were presumed benign and followed.
The initial clinical stages of disease in the eligible patients (n = 303) after standard staging procedures were the following: IA in 125 patients (41.2%), IB in 106 patients (35%), IIA in 5 patients (1.7%), IIB in 18 patients (6%), and IIIA in 49 patients (16.2%). PET correctly identified M1 disease (presumed to be M0 by standard imaging procedures) in 15 of 287 patients (5.2%). The sites of metastases subsequently confirmed by biopsy or radiographic study were bone in 5 patients, adrenal gland in 2 patients, extrathoracic lymph node in 2 patients, lung in 1 patient, liver in 1 patient, and colon in 1 patient. PET identified metastatic disease at multiple sites in 3 subjects, and therefore these patients did not undergo surgical intervention. There were 3 additional patients who had M1 disease discovered at thoracotomy that was missed.
by both standard imaging and PET (2 patients with lung disease and 1 patient with pleural disease). PET also identified 3 patients with second primaries: colon, lung, and thyroid. In 19 of 287 evaluable patients (6.6%), possible M1 disease indicated by PET was subsequently excluded by additional radiographic studies (13), by biopsies (4), or at the time of surgical resection (2). The false-positive foci were in brain (8), distant lymph nodes (2), bone (2), rectum, thyroid, kidney, larynx, esophagus, liver, and contralateral lung. There were 8 additional patients with M1 disease indicated by PET; these findings were considered to be false positive in the judgment of the surgeon, and no confirmatory studies were performed. Three of the patients with presumed false-positive PET findings for M1 disease died within 6 months, indicating that PET may have been correct. The sensitivity of PET for M1 disease was 83%, and the specificity was 90%. The negative predictive value (NPV) was 99%, and the positive predictive value (PPV) was 56%. The ability of CT to predict nodal disease is 37%, but the NPV and PPV of CT were 81% and 58%, respectively. Of 62 eligible patients with positive findings on history review or physical examinations, 4 had M1 disease (of whom only 1 had symptoms consistent with the M1 site).

The ability to correctly classify nodal disease (N0, N1, and N2/N3) confirmed pathologically by CT alone, PET alone, and CT combined with PET is shown in Table 2. Of the 303 eligible patients, 37 underwent a noncontrast chest CT; 25 of these patients were evaluated for isolated peripheral nodules, and 12 patients did not receive contrast agent because of allergy, patient refusal, or other reason. PET correctly identified 128 of 168 patients (76%) with N0 disease, whereas CT correctly identified 147 of 168 patients (88%) with N0 disease (P > .5). PET was superior to CT in the identification of N1 disease (42% vs 13%; P = .0177). PET correctly detected mediastinal lymph node disease (N2/N3) in 58% of cases compared with 32% for CT (P = .0041). The combined PET and CT readings did not significantly alter the findings of PET alone.

The ability of PET and CT to predict mediastinal nodal disease is shown in Tables 3 and 4. For all eligible patients, the sensitivity of PET for N2 and N3 disease was 61% and the specificity was 84%. The NPV of PET was 87%, and the PPV was 56%. The ability of CT to predict nodal disease is shown in Table 4. For all eligible patients, the sensitivity of CT to detect N2 and N3 disease was only 37%, but the specificity was 91%. The NPV and PPV of CT were 81% and 58%, respectively.

CT identified 48 evaluable patients with N2 disease considered resectable by the participating surgeon. PET correctly upstaged the disease in 5 of these patients to N3 or M1. PET correctly identified unsuspected IIIA, IIIB, or M1 disease considered inoperable in another 38 patients. Therefore in at least 43 evaluable patients, nontherapeutic thoracotomy could be avoided using information gained from PET after standard staging procedures. When patients with benign lesions correctly identified by PET and who underwent thoracotomy are included (n = 18), nontherapeutic thoracotomy could potentially be avoided in 61 of 303 patients (20.1%). These results validate the 2-stage design of this trial, such that the uniform minimum variance unbiased estimate for the probability of FDG-PET contraindicating surgery was 0.201 (95% Jennison-Turnbull confidence interval 0.158-0.251).7,8

Discussion
From the patient’s viewpoint, staging of NSCLC is most critical in assessing candidacy for potentially curative surgical resection. The discovery of extrathoracic metastatic

<table>
<thead>
<tr>
<th>TABLE 2. Correct classification of nodal status in patients with primary lung cancer by imaging modality*</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Confirmed pathology</strong></td>
</tr>
<tr>
<td>------------------------</td>
</tr>
<tr>
<td>N0</td>
</tr>
<tr>
<td>N1</td>
</tr>
<tr>
<td>N2/N3</td>
</tr>
</tbody>
</table>

CT, Computed tomography; PET, position emission tomography.

*Fraction of patients with nodal status confirmed pathologically and correctly staged by the imaging modality.

<table>
<thead>
<tr>
<th>TABLE 3. Lymph node status comparing PET with final stage (all eligible patients; n = 302)*</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>By PET</strong></td>
</tr>
<tr>
<td>-----------</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>N0/N1</td>
</tr>
<tr>
<td>N2/N3</td>
</tr>
<tr>
<td>Total</td>
</tr>
</tbody>
</table>

PPV, Positive predictive value; NPV, negative predictive value.

*Final stage equals nodal stage as determined radiographically or pathologically, including patients who did not have nodal status confirmed with biopsy, mediastinoscopy, or surgical resection. For N2/N3 disease: sensitivity = 61%; specificity = 94%; PPV = 56%; NPV = 87%.

<table>
<thead>
<tr>
<th>TABLE 4. Lymph node status comparing CT with final stage (all eligible patients; n = 302)*</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>By CT</strong></td>
</tr>
<tr>
<td>-----------</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>N0/N1</td>
</tr>
<tr>
<td>N2/N3</td>
</tr>
<tr>
<td>Total</td>
</tr>
</tbody>
</table>

*Final stage equals nodal stage as determined radiographically or pathologically, including patients who did not have nodal status confirmed with biopsy, mediastinoscopy, or surgical resection. For N2/N3 disease: sensitivity = 37%; specificity = 91%; PPV = 58%; NPV = 81%.
disease is a contraindication to surgery except in very specialized circumstances (eg, solitary brain metastasis). Patients with involvement of mediastinal lymph nodes, either ipsilateral (N2) or contralateral (N3), are usually not considered for primary surgical treatment. Some surgeons consider selected patients with single-station N2 involvement operable. The Z0050 protocol allowed inclusion of these patients to mirror clinical practice. Those patients without distant metastasis and metastatic lymph nodes (N0) or with only intrapulmonary or hilar lymph node disease (N1) are generally considered suitable to undergo surgical resection.

PET adds a new dimension to the anatomic staging tools routinely used to evaluate NSCLC. Because of relatively poor spatial resolution of PET, the functional information it provides is considered complementary to the anatomic information provided by CT. This complementary use is especially important in the accurate assessment of the tumor itself (T status) and the evaluation of intrapulmonary nodes close to the mediastinum. Our study shows that although neither CT nor PET is highly successful in identifying N1 disease, PET is significantly better.

For mediastinal lymph node staging, PET has been found to be superior by meta-analysis in detecting and excluding metastatic disease when compared with CT.9,10 For all prospective studies comparing PET and CT in at least 50 patients, the overall accuracy of FGD-PET in the detection of mediastinal lymph node involvement was 90% with a sensitivity of 85% and a specificity of 92%.11 CT was accurate in only 65% of cases with a sensitivity of 75% and a specificity of 80%. In the present study, PET was significantly better than CT in the correct identification of N2 or N3 disease, but the sensitivity of PET for mediastinal disease (61%) was considerably less than in previous reports. The result, however, is very similar to the sensitivity in a prospective series reported by Gonzalez-Stawinski and colleagues from Duke.12 The sensitivity of PET for mediastinal staging in recent studies is documented in Table 5.12-16

| Author                          | Year | No. patients | Sensitivity (%) | Specificity (%) |
|---------------------------------|------|--------------|----------------|----------------|----------------|
| Roberts et al13                 | 2000 | 100          | 88             | 91             |
| Gupta et al14                   | 2001 | 77           | 87             | 91             |
| Kerstine et al15                | 2002 | 237          | 82             | 82             |
| Vesselle et al6                 | 2002 | 118          | 81             | 96             |
| Gonzalez-Stawinski et al12     | 2003 | 203          | 64             | 77             |
| Z0050                          | 2003 | 271          | 61             | 84             |

that if the PET finding was negative in N2 or N3 stations, there was a 1% to 8% likelihood of mediastinal cancer. The authors believed the findings justified the continued use of mediastinoscopy. In the study by Vesselle and colleagues,16 PET underestimated the presence of N2 nodal disease in 6.8% of patients. Graeter and colleagues17 analyzed the results of 82 patients with NSCLC with radiologically suspected mediastinal lymph nodes and found the NPV of PET to be 98.4%. They therefore concluded that mediastinoscopy could be omitted in this situation. However, this study is not comparable to most others, in which the findings of PET were evaluated in patients with negative CT scans. The Z0050 trial validates that mediastinoscopy must still be considered the gold standard. Many reports substantiate that PET can be very useful in guiding mediastinal biopsy, particularly when PET demonstrates a lesion in a nodal station not amenable to mediastinoscopy.

Discovery of distant metastatic disease precludes surgical resection. Routine staging includes bone scintigraphy, CT or MRI of the brain, abdominal CT, and occasionally ultrasonography. FDG-PET is an attractive modality for finding distant metastases because the whole body can be evaluated in 1 imaging session. However, FDG-PET has not been useful for the detection of brain metastasis because the lesions are often obscured by the high FDG uptake in normal gray matter.18 Although bone scintigraphy is reasonably sensitive, it is hampered by its low specificity requiring additional studies. FDG-PET was found superior to bone scintigraphy in 2 studies.19,20 PET may also identify earlier disease, such as narrow metastases. Approximately 20% of patients with NSCLC present with adrenal masses, but many are benign adrenal adenomas. The usefulness of FDG-PET in the evaluation of an adrenal mass has been reported, and 3 studies found a sensitivity of 100%.21-23 Studies evaluating the performance of FDG-PET for detecting liver metastases from NSCLC have been limited. Isolated liver metastasis is rare and usually identified by CT or ultrasonography, but PET may be more specific. Thus, FDG-PET may be the most appropriate modality in combination with CT or MRI of the brain to assess extrathoracic disease.

FDG-PET has been reported to detect occult extrathoracic metastases in 11% to 17% of patients with tumors deemed resectable by conventional methods.4,14,24,25 In this study of potentially resectable tumors as determined by mandated complete staging procedures, the detection of unsuspected metastases was much lower at 6.3%. However, the study was not designed to assess the utility of PET as an initial evaluation for metastatic disease, because patients with metastases discovered by routine scans were not registered. PET identified 19 of 287 patients (6.6%) with findings indicative of M1 lesions that were subsequently proven not to be metastatic. Of importance is the fact that a signif-
significant number of patients with M1 disease identified by PET could have been denied curative surgery. In 3 patients, new primary cancers were discovered. In another 8 patients, abnormal uptake in the brain was noted. Further assessment is required when M1 disease is identified by PET, especially if it is in a single site or in the brain.

The primary objective of this trial was to evaluate the usefulness of adding PET for staging disease in patients with documented NSCLC or believed to have bronchogenic carcinoma and who had undergone complete routine staging procedures. This trial therefore differs from most other series in which incomplete staging was performed. The ACOSOG Z0050 trial confirms the usefulness of PET in this setting. Advanced, unresectable disease was found in 43 of 303 patients (14.2%). When the identification of benign disease is also considered, approximately 1 in 5 patients could avoid nontherapeutic thoracotomy (20.1%). This study validates the findings of the PLUS trial in which addition of PET to conventional workup prevented unnecessary surgery in 1 of 5 patients with suspected or proven NSCLC.

The Z0050 study differs from the PLUS trial by requiring radiographic staging in all patients before PET. The conventional workup in the PET cohort of the PLUS trial included bone scintigraphy in 27%, brain CT and MRI in 5%, and abdominal CT and ultrasonography in 46%. The PLUS trial also included twice as many patients with locally advanced (IIIA and IIB) disease compared with the Z0050 trial.

Although PET is better than CT in identifying mediastinal disease, a significant number of patients will still have N2 or N3 disease when PET findings are negative. PET positivity is dependent on tumor cell volume, tumor density, and avidity for FDG. In many cases, lymph node positivity corresponds to microscopic deposits of tumor cells. Invasive mediastinal staging therefore is still recommended. A high false-positive rate is found in geographic areas where granulomatous or other inflammatory diseases (eg, histoplasmosis) are common. Positive PET findings in the mediastinum should be confirmed, and PET may be very useful to direct the location and type of invasive staging. In patients with disease well staged by routine radiologic imaging, PET found only an additional 6.3% of patients with M1 disease. In view of the low specificity of bone scintigraphy for osseous metastasis and CT scans for adrenal metastasis, PET may be more useful than indicated by this study. M1 disease identified by PET requires further evaluation because 6.6% of patients in the Z0050 trial potentially would have been denied surgery if treatment decisions were based on PET findings alone.

There were several limitations to this study. Patients with suspected, but not biopsy-proven, NSCLC were eligible for the study. A total of 11.6% of patients had benign disease. However, the inclusion of such patients reflects clinical practice. As noted, discrimination of N2 and N3 disease was not always possible because of the clinical practice of halting mediastinoscopy when biopsy at 1 level is positive. Patients with stage IIIA disease were allowed in the study if the surgeon considered the disease resectable after routine staging. A final weakness was the lack of required follow-up in patients not undergoing resection.

The usefulness of PET in the staging of lung cancer is evolving. The Z0050 trial has shown that currently PET has the ability to prevent nontherapeutic thoracotomy in a high percentage of patients, even after thorough staging procedures. The initial enthusiasm for PET as the sole staging tool for mediastinal disease must clearly be tempered. The ability of PET-CT fusion imaging to improve on the sensitivity and specificity of PET alone has recently been reported, but assessment of the role of this new approach awaits further study.

References

14. Gupta NC, Tamim WJ, Graeber GG, Bishop HA, Hobbs GR. Medi-


**Appendix 1**

Z0050 Thoracic Physician Group: Bryan Meyers (St. Louis, Mo), Valerie Rusch (New York, NY), Kemp Kernsteine (Iowa City, Iowa), James Luketich (Pittsburgh, Pa), Gary Chmielewski (Royal Oak, Mich), John Niederer (Madison, Wis), Neri Cohen (Richmond, Va), Michael Collins (Salt Lake City, Utah), Marshall Urist (Birmingham, Ala), Kenneth Tanabe (Boston, Mass), Thomas D’Amico (Durham, NC), Carolyn Reed (Charleston, SC), Peter Blumencranz (Clearwater, Fla), Mark Ferguson (Baltimore, Md), Jemi Olak (Park Ridge, Ill), Robert Cameron (Los Angeles, Calif), Kenneth Kesler (Indianapolis, Ind), Mark Krasna (Baltimore, Md), Paul Waters (New York, NY), David Jablons (San Francisco, Calif), Jack B. Thigpen (Lakeland, Fla), and Robert Kennan (Pittsburgh, Pa).

**Discussion**

Dr Robert J. Cerfolio (Birmingham, Ala) Dr Reed, that was an excellent presentation concerning a very important clinical topic. Perhaps the most important message that can be gleaned from this project is that multiple surgeons along with radiologists and pathologists from many different institutions can construct and then carry out a high-quality study in an attempt to answer an important clinical question that we all have. It takes an incredible amount of time, energy, and work from a whole lot of people, including multiple institutional review boards, and I hope we all continue to work in this manner. I want to first congratulate you on completing and now presenting this trial.

Now, as general thoracic surgeons, we are always trying to improve our patient selection for those afflicted with lung cancer. There is no question that PET with FDG allows us to do this, but unfortunately it is not as great as we all hoped. Although PET is good, it “ain’t the great hope” that we all envisioned. It clearly helps us to target areas to biopsy, but there are lots of false-positives and negatives. Now, many hospitals claim to have a PET when indeed they have a pseudo-PET, and I believe this is one of the biggest problems facing PET scans today. Many patients come to see us in our clinic and already have had a PET performed, oftentimes at an unknown center, on unspecified cameras.

So my first question to you is, what tricks do you have that we can use in our clinics that can help us determine what type of machine was used to perform the PET scan and who read it, which may be even more important. How can we help educate insurance companies to pay for a repeat PET on a real dedicated PET with the proper thickness camera when it is indicated?

The second problem facing PETs today is the SUVs. This acronym, which stands for standardized uptake values, is really an oxymoron more than an acronym. SUVs are clearly not standardized. An SUV of 3.0 at my center, which may mean a nodule is malignant, may be an SUV of 2.2 at your center, meaning it might be benign. What steps are being done to help with this problem?

That brings us to PET for the mediastinal lymph nodes. In your report you chose not to report the sensitivity, specificity, PPV, NPV, or accuracy of FDG-PET for N1, N2, and N3 nodes separately, as we and most others have reported. Moreover, you grouped the N2s and N3s together. The results you reported are lower than other reports. Can you explain why? Did you see many false-positives from histoplasmosis and silicosis as we and others have reported? False positivity in the lymph nodes is a real problem with PET, and we believe a positive PET finding for a mediastinal lymph node only means that a biopsy is indicated. Do you agree? We have shown that PET seems to be more accurate in some N2 stations than others. Did you evaluate the accuracy in different lymph node stations? And although false-positives are more common than false-negatives, it is the latter that I really worry about. Can you discuss your false-negatives? Were they from patients with microscopic disease? I know you mentioned bronchoalveolar carcinoma in one of your slides, but I did not see anything about carcinoid. Can you tell us about that?

My fourth question concerns PET for M1 disease—it seems to be quite good. You have shown that PET is very accurate at evaluating the liver and adrenal glands. However, what about the brain (if scanned by PET) and the bone? We all agree that if a PET indicates a brain lesion, an MRI of the brain is indicated, but do you currently recommend an MRI for all asymptomatic patients? Is PET often falsely negative for the brain? And how do you handle a PET that indicates M1 disease in the hip, for example? Do you obtain a bone scan or do you go right to an MRI, and if the MRI is positive for the bone, is that study in and of itself good?
enough without tissue to deny a patient a pulmonary resection? We have had several false-positive MRI scans of the bone.

Finally, you have shown that PET and CT are poor for N1 disease, and again your reported accuracy is lower for both than in many other reported series. Can you explain why? And do you have any experience with PET-CT fusion for N1 or N2 disease? We have completed a prospective randomized trial that compared PET with PET-CT fusion and found that PET-CT is superior for detecting N1 disease. There are many ongoing trials evaluating the use of chemotherapy for early lung cancer, and the presence of N1 disease may change our preoperative management. Therefore, this issue may turn out to be clinically important. Would you recommend to someone who is about to build a PET center to buy a PET-CT fusion machine or just a PET?

We are indebted for all your hard work. This is an important article and a great presentation, and I thank you and the association for your time.

**Dr Reed.** I thank Dr Cerfolio for his remarks and thoughtful questions. This is the first ACOSOG trial to be completed. We have learned much, and clearly we still need to educate surgeons about clinical trials. I hope that future trials will be better, and I am sure they will.

Dr Cerfolio’s questions about PET scanners and the SUVs recorded from different institutions are important points. I do not have specific answers for him. In this trial the PET scanning protocol was very detailed, and studies from each participating institution were reviewed and had to be approved by the ACOSOG PET Quality Assurance Committee. And you alluded to something, I think, that may be more important: A dedicated, consistent radiologist with whom the surgeon interacts is probably more important than the scanner.

One of the weaknesses of this study was our inability to accurately separate N2 and N3 disease. When PET findings were positive for multiple levels, the surgeon often performed a mediastinoscopy to verify positivity and stopped after one of the levels was positive by frozen section. I think your point about N2 levels is important. We did not look at the different N2 stations, but we could go back and do that and should. I can tell you that in a little over half of the cases, positivity at pathology was only for a single level. We obviously, as you do, believe that a PET positive finding on the scan should be verified pathologically.

The lower sensitivity of PET than previously reported for mediastinal disease is really not surprising, and I think more and more surgeons are seeing this. PET positivity is dependent on tumor cell volume, tumor density, and avidity for FDG. In many cases, lymph node positivity corresponds to microscopic deposits of tumor cells in these lymph nodes.

One of the lessons we learned from this trial is that a single site of PET positivity must be confirmed. PET is very sensitive for bone metastases, as has been documented in the literature. I think most of us would consider a positive MRI after a PET scan as a confirmatory procedure. I will tell you that of the many false-positive PET findings for M1 disease, 6 were actually in the brain. As always, however, the onus is on us to prove that a patient’s tumor is not resectable.

**Dr Frank C. Detterbeck (Chapel Hill, NC).** There is a dramatic difference between the incidence of picking up distant metastatic disease after a negative clinical evaluation depending on whether the patient has clinical stage I, II, or III disease. Old data of head CT and bone scan showed that 5% of patients with clinical stage I disease had positive findings. There are data for PET and patients with clinical stage I disease showing a less than 10% incidence of systemic metastases, whereas patients with stage III disease have a 20% or 30% incidence.

Have you looked at your results by clinical stage? And a related question is, what is the bottom line? Do you recommend PET for all patients with lung cancer or for patients with clinical stage I, II, or III disease? What is it?

**Dr Reed.** First of all, concerning your first question on clinical stage, of those patients with IIIA disease clinically (eg, there were, I believe, 42 of them), PET findings upstaged the disease in 12 cases. So indeed, as the stage goes up, PET will show you increased positivity. What is the bottom line? I think Bob alluded to the fact that we had all hoped PET would be the magic tool, the single tool, and clearly it is not. I think most of us are now using PET more for extrathoracic metastases, particularly bone, adrenal, and liver. We all appreciate that PET is not useful for the brain. I think most of us have gone back to using more mediastinoscopy because PET has not been as sensitive in the mediastinum. So I think at a lot of institutions now the protocol is to add the PET after a chest CT to look for that extrathoracic metastases, and whether you add the MRI or head CT may be very dependent on the clinical examination. The NPV of 87%, everybody has their own threshold, as you know. It depends on your institution and what the PET findings are in your institution, your personal threshold, and other articles in the literature.