Comparison of Whole-body FDG-PET to Bone Scan for Detection of Bone Metastases in Patients with a New Diagnosis of Non-Small Cell Lung Cancer

Sendhil Kumar Cheran

Russell Harris, M.D., M.P.H.
Advisor, Health Care and Prevention
Public Health Leadership Program
School of Public Health
University of North Carolina at Chapel Hill

Edward F. Patz, M.D.
Professor, Department of Radiology
Duke University Medical Center
Comparison of Whole-body FDG-PET to Bone Scan for Detection of Bone Metastases in Patients with a New Diagnosis of Non-Small Cell Lung Cancer

Sendhil Kumar Cheran

ABSTRACT

Objective. In patients with lung cancer, whole-body 2-deoxy-2-\(^{18}\)F]fluoro-D-glucose with positron-emission tomography (FDG-PET) scan and bone scintigraphy are common staging studies, both with the ability to detect osseous metastasis. There is some evidence that PET scan is superior to bone scan and that bone scan can be replaced by PET scan. The purpose of this study was to compare the test characteristics and agreement of FDG-PET to bone scintigraphy for the detection of bony metastases in the staging of patients with newly diagnosed lung cancer.

Subjects and Methods. We queried the tumor registry and nuclear medicine database at Duke University Medical Center to identify all patients between July 1998 and August 2002 with a new diagnosis of lung cancer, FDG-PET scan, and a bone scan prior to therapy. We retrospectively reviewed these patients’ radiologic reports and entered them into a database. We used specific clinical criteria to confirm bone metastases, integrating all available clinical information. We then calculated the sensitivity, specificity, and likelihood ratios for each study.

Results. Two hundred and fifty-seven patients fulfilled the entrance criteria. One hundred and four patients (40%) presented with stage IV disease, and we confirmed bone metastases in 57 (22%) patients. We calculated the sensitivity values of PET and bone scan as 91% and 63-75%, and specificity values as 96%
and 92-95%, respectively. Likelihood ratios of positive PET and bone scans were greater than 20. Likelihood ratios of negative PET and bone scans were 0.08 and 0.16, respectively. The likelihood ratio of equivocal bone scans was 3.5. The weighted-kappa statistic suggested moderate agreement between the two modalities ($\kappa_w = 0.510$, 95% CI, 0.402 – 0.618).

Elimination of bone scan in this cohort would have resulted in 7 (3%) of 257 patients with false positive PET results for bone metastasis, but in no case would patient management have been adversely affected. Similarly, 5 (1%) of 257 patients with bone metastases had false negative PET results, but in all cases there was other radiographic evidence of extrathoracic metastasis and Stage IV disease by CT scan or PET.

**Conclusions.** The use of both whole-body PET and bone scintigraphy as initial staging studies in lung cancer patients may provide redundant information about the presence of bony metastases. The improvement in accuracy and sensitivity with PET suggests bone scan could potentially be eliminated from the staging evaluation at presentation, but these results are limited by sub-optimal study design. We recommend a prospective trial with appropriate verification of bony metastases to confirm these results.
INTRODUCTION

Background. Lung cancer is the leading cause of cancer death in the world.¹ In the United States alone, lung cancer will cause an estimated 157,000 deaths this year, equal to the number of deaths from prostate, breast, and colon cancer combined.²

Of the roughly 153,000 Americans who die each year from lung cancer, smoking accounts for 125,000 (82%) deaths.³ Smoking cessation has long been known the most effective form of lung cancer prevention. There is a dose-response relationship between the amount of tobacco exposure and the development of lung cancer.¹ Increased public awareness of the numerous health risks of smoking including lung cancer resulted in an overall 40% decline in smoking prevalence in the United States between 1965 and 1990. Since then, smoking prevalence has remained constant. Currently, 25.7% of adult US males (24.3 million) and 21.5% of adult US females (22.2 million) are smokers. The percentage of high school children who report current cigarette smoking is even higher at 28.5%. Projections of current smoking patterns suggest lung cancer will remain the world’s leading cause of cancer death for at least the next several decades to come.⁴

Staging of Lung Cancer. Among patients who develop lung cancer, accurate staging is essential for patient management and in determining the optimal therapeutic strategy.⁵-⁷ Identification of patients with distant metastasis (Stage IV disease) precludes curative surgery. Patients staged with localized Stage I or II disease have greater than 50% 5-year survival, but this figure drops
sharply to 20% in patients with Stage III disease and less than 2% 5-year survival in patients with Stage IV lung cancer.8-10 The aforementioned statistics are subject to lead-time bias and thus unreliable for determining lung cancer screening policy. However, the data show clearly that Stage IV lung cancer patients have a dismal prognosis and confirm the current view that Stage IV patients cannot be cured by surgical resection. Unfortunately, only 15% of lung cancer patients present with localized Stage I or Stage II disease.5

Inability to detect Stage IV disease in lung cancer patients on presentation results in higher cost and increased morbidity associated with futile surgical intervention. Therefore, diagnostic workup of lung cancer patients always starts with a search for distant metastasis.11,12 A series of radiographic tests performed upon initial diagnosis of lung cancer is used to identify patients with Stage IV disease. Clinical staging at the time of presentation is typically performed using a series of targeted diagnostic imaging studies including computed-tomography (CT) of the thorax through the liver and adrenal glands, CT and/or magnetic resonance (MR) imaging of the head, and radionuclide bone scintigraphy (bone scan). In the case of solitary lesions of unknown origin in liver, adrenal gland, or bone, biopsy is occasionally used to remove diagnostic uncertainty.

An important consideration is that 40% of patients undergoing “curative” resection have a recurrence of lung cancer.8 This fact suggests current methods of staging are frequently missing the presence of distant metastasis. Thus, researchers have been searching for new modalities to more accurately stage lung cancer patients and detect regional and distant spread.
FDG-PET in Lung Cancer. More recently, researchers have suggested whole-body 2-deoxy-2-[\(^{18}\text{F}\)]fluoro-D-glucose (FDG) with positron-emission tomography (PET) as an alternative study, which could provide a more accurate and efficient diagnostic approach than the combination of conventional imaging studies.\(^{13-17}\) FDG-PET is a type of functional imaging based on the principle that malignant tumors have a higher metabolic rate than benign tumors or background physiologic activity. Malignant cells thus preferentially accumulate higher concentrations of FDG than surrounding tissue. FDG is irreversibly trapped in cells and unable to undergo further metabolism after the first phosphorylation step of glycolysis, which allows for the detection of areas of high metabolic activity corresponding to malignancy or infection.\(^{18}\) False-positive PET results occur in areas of increased metabolism such as infection. Another potential drawback of PET scanning is the inability to localize areas of hypermetabolism with anatomic precision.

PET scanning has become an accepted tool for diagnostic evaluation and initial staging of lung cancer as evidenced by approval of Medicare coverage since January 1998.\(^{13,19}\) Medicare expanded coverage to the diagnosis, staging, restaging of non-small cell lung cancer in July 2001.\(^{19}\)

A recent meta-analysis of good quality examined the overall accuracy of PET for regional lymph node staging in lung cancer, showing it to be superior to that of conventional staging with CT.\(^{20}\) While none of the 29 included studies fulfilled all of the authors’ prospectively-defined criteria for study quality, all studies used an objective gold standard for evaluation of lymph nodes. The gold
standard was pathologic evaluation of lymph nodes from either mediastinoscopy, thoracotomy, or autopsy. Seventy percent of the studies in the meta-analysis were prospective.

Another meta-analysis of lesser quality analyzed PET accuracy in the evaluation of malignancy in indeterminate pulmonary nodules. This meta-analysis placed the sensitivity and specificity of PET at 97% and 78%, respectively.21 Many of the studies had incomplete blinding, and the authors assessed the included studies as being of fair quality overall. The authors defined minimum eligibility criteria for inclusion in the meta-analysis and later performed an assessment of quality of the included studies. In other words, study quality was not part of the inclusion criteria. As a result, 5 of 37 FDG-PET studies included in the meta-analysis did not even meet the authors' quality criteria for having an appropriate reference standard.

**FDG-PET in Extrathoracic Staging of Lung Cancer.** There is controversy in the medical literature about the ideal imaging approach to search for extra-thoracic metastases in bone and other distant sites. At Duke University Medical Center, the pattern of diagnostic imaging in patients with lung cancer differs depending on the type of clinic. All patients presenting to thoracic surgeons at DUMC for diagnostic workup and possible curative surgical resection of lung cancer undergo thoracic CT scan and whole-body PET scan. Thoracic surgeons at Duke do not search for lung cancer metastasis to bone or brain unless patients demonstrate clinical symptoms such as bone pain or focal neurological symptoms.5,6,14 However, at the Duke thoracic oncology clinic, all patients
determined to be Stage IIIB, Stage IV, or medically-inoperable invariably undergo bone scan and imaging of the head.\textsuperscript{11,14} The frequency and pattern of utilization of PET in the Duke thoracic oncology clinic depends on the pattern of use of previous imaging modalities and corresponding results. While patients in the oncology clinic often have metastatic disease that rules them out from curative surgery, identification of bone metastases can provide important information about baseline level of disease and response to therapy.

Several studies including 2 prospective cohort studies and 1 randomized, controlled trial have shown PET to be beneficial for extra-thoracic staging in the overall search for metastatic disease in lung cancer.\textsuperscript{14,17,21-23} In 11% of patients, PET can identify clinically silent metastasis missed by other conventional imaging modalities.\textsuperscript{23} The results of PET change patient management in greater than 26% of lung cancer patients by upstaging them to palliative therapy from curative surgery.\textsuperscript{22-24}

**Bone metastases.** Bone is a common site of metastasis in patients with lung cancer, but its diagnosis can be difficult.\textsuperscript{14} The pathophysiology of bone metastasis is thought to begin with seeding of bone by micrometastases after spread of malignant tumor cells into the bloodstream.\textsuperscript{25} Micrometastatic spread to bone can be detected by immunohistochemical stains of bone marrow aspirates.\textsuperscript{26} However, the simple presence of micrometastatic lung cancer in bone is not necessarily diagnostic of clinically important metastasis to bone, nor is it a good indicator of patient outcome.\textsuperscript{27} Many clinicians exclude the diagnosis of bone metastasis in lung cancer patients without clinical symptoms, because the
negative predictive value of absent bone pain is greater than 90% according to some studies.\(^{14}\) This approach is fraught with uncertainty, since some series report 40% of lung cancer patients with proven bone metastasis as asymptomatic.\(^{28}\)

**Bone Scintigraphy in Lung Cancer.** To date, the most common imaging study used to detect gross metastatic lesions of bone is bone scan with technetium-99m labeled methylene diphosphate. Other second-line imaging tests are MR imaging of bone and plain film radiography. Unlike FDG-PET, which detects increased glucose metabolism, bone scan is an indicator of bone turnover through incorporation of radiolabelled phosphate into bone. Bone scan is an imperfect imaging modality for detection of bone metastasis in lung cancer. It works best at detecting sclerotic metastases, because they contain areas of increased bone turnover and bone deposition. But lung cancer sometimes results in lytic metastases, resulting in a decreased sensitivity of bone scan.\(^{29}\) Sensitivity of bone scan is quite high according to previous studies, but the specificity is low with frequent false positives.\(^{25,28}\) This is largely due to radiotracer uptake in a multiplicity of benign conditions such as arthritic joints, previous skeletal trauma, Paget’s disease, or benign inflammatory processes in a pattern indistinguishable from metastasis.

**Bone Scan and FDG-PET in Detection of Lung Cancer Bone Metastases.** Given the ability of PET to detect osseous metastasis, several small series have indicated that the high accuracy of PET may obviate the need for a bone scan.\(^{17,28,30-32}\) Only four published studies have examined closely the
relative accuracy of PET scan and bone scan in the detection of bone metastases in lung cancer patients.\textsuperscript{17, 28, 30, 31} We identified 3 studies through a search of the Medline and CINAHL databases with the search terms “bone scan”, “bone scintigraphy”, and “positron emission tomography” and the MeSH term “Lung Neoplasm”.\textsuperscript{17, 28, 30} We obtained one study through consultation with an expert in the field.\textsuperscript{31}

No study made routine use of a gold standard such as bone biopsy. All studies employed a reference standard of follow-up bone scans over a period of several months, with occasional use of other radiographic imaging or bone biopsy when deemed appropriate by the treating physician. The Marom et.al. study had the most variable follow-up, ranging from 1 to 8 months, whereas all other studies had a minimum clinical follow-up period of 9 months.

Sensitivity and specificity values ranged from 90-93\% and 98-99\%, respectively. Sensitivity and specificity values of bone scan were much more variable, ranging from 50-94\% and 61-92\%, respectively. The wide range is due in part to study design. Study populations ranged from 48 to 110 patients but included no more than 21 patients determined in the final analysis to have osseous metastasis. The small number of patients with bone metastases resulted in wide confidence intervals of all results.

All studies utilized blinded interpretation of radiographs by at least 2 experienced reviewers. Of the four cohort studies, Marom et. al. was the only prospective one. No study accounted for the effect of chemotherapy on the potential interval resolution of metastatic bone lesions.
Bury et al. viewed any abnormal uptake on bone scan as representative metastatic disease unless there was clear evidence of degenerative disease by prior radiographs or high clinical suspicion of the treating physician. This stringent criterion maximized the sensitivity of bone scan but handicapped its specificity. It is also not representative of bone scan interpretation by nuclear medicine physicians at DUMC or most other institutions.

Hsia et al. compared PET and bone scan on a lesion-by-lesion basis, whereas the other studies performed a patient-by-patient analysis. Thus, it was impossible to determine the effect on individual patients of excluding bone scan or PET scan.

Despite the consistency of results, all four studies had shortcomings in design. No study had a solid reference standard, and all studies had a very small number of patients affected by bone metastasis.

Equivocal or Indeterminate Results. Since bone scan is well-known to demonstrate lesions of variable uptake which cannot differentiate between metastatic and degenerative disease, these lesions cannot be described as clearly positive or negative. Thus, results of bone scan are not dichotomous. Similarly, there can be difficulty differentiating on PET between location of tumor in peripheral lung or adjacent rib. However, the aforementioned studies have reported all PET and bone scan results as either positive or negative for bone metastasis, ignoring the clinical reality of equivocal or indeterminate results. This methodology of artificially forcing indeterminate diagnostic results into firm diagnostic categories of positive or negative raises questions about the validity
and clinical relevance of subsequent analysis and conclusions. Resultant calculations of sensitivity, specificity, and other diagnostic test characteristics can be misleading or inaccurate, with unintended financial or diagnostic implications. There is no consensus in the epidemiology literature of how to analyze indeterminate or equivocal results of diagnostic tests. At the very least, investigators should report the existence of this diagnostic category and clearly specify their chosen method of addressing the category in the analysis. In general, indeterminate or equivocal results should not be excluded from the analysis, nor should they be placed in the standard 2x2 matrix. Researchers should consider developing likelihood ratios for the intermediate results.

**Future of FDG-PET in Lung Cancer.** At DUMC, the cost of bone scan is approximately $400-$500, whereas PET scan is approximately $2200-2500. Although PET studies are clearly more expensive than bone scan and CT scan, utilization of PET is spreading beyond the realm of the tertiary-care, academic medical center. As the body of evidence supporting PET utilization in staging and diagnosis of lung cancer and other less-common cancers grows, availability of PET will increase rapidly in community medical practices.

A recent prospective cohort study of 50 patients showed a statistically significant improvement in diagnostic accuracy with fusion of PET and CT images by combining the benefits of PET’s functional imaging capabilities with the anatomic specificity of CT scan. The gold standard was histopathological assessment of tumor stage and regional lymph nodes, and histopathological or radiographic assessment of distant metastasis. As a result of this study, many
physicians who treat lung cancer now view PET-CT fusion as a new standard of care for lung cancer staging and re-staging, and the current trend in radiology is toward the purchase of combination PET-CT fusion cameras. Since PET-CT fusion provides superior results compared to either modality alone, it is highly likely that a fused PET scan will accompany all future thoracic CT scans in lung cancer patients. The push for rapid acceptance and adoption of this new technology is understandable given the high prevalence of lung cancer and its dismal prognosis. However, immediately shifting the bulk of staging radiography to PET-CT fusion scans seems rash, given the paucity of supporting evidence and the high equipment cost.

One current issue in extrathoracic staging of lung cancer is whether bone scan can provide clinical information over and above that supplied by PET. Evaluation of this question must begin with a comparison of the diagnostic test characteristics of both studies. The purpose of this study was to compare the agreement and test characteristics of PET scan to bone scan for the detection of distant bone metastases in patients with newly diagnosed lung cancer to determine if these were complementary or redundant studies. Based on the work of previous authors, we expect FDG-PET to have a statistically significant increase in sensitivity and specificity over bone scan.
SUBJECTS AND METHODS

Study Design

Several authors have described criteria to evaluate the quality of studies of diagnostic tests.\textsuperscript{36,37} Two of the most important are blinding and use of a gold standard. In the design of diagnostic imaging tests, blinding of test interpreters is important because of the inherent subjectivity of evaluating the images. Outside clinical information is highly likely to influence the interpretations of unblinded radiologists. Because of the retrospective design, we could not blind the radiologists interpreting the bone or PET scans in this study. Due to the large number of scans, we were able to recruit only one blinded radiologist to provide blinded interpretations, but these were only for scans with ambiguous or equivocal reports. Ideally, two blinded radiologists would provide independent interpretations of all scans with disputes resolved by a third observer or consensus. Their interobserver agreement would be calculated in the form of a $\kappa$ statistic.

The gold standard for bone metastasis is a positive bone biopsy in a region of bone pain or positive radiographic findings. This study employed an alternate but imperfect reference standard. The reference standard in this study and previous studies incorporated the results of PET scan and bone scan to determine the presence of bone metastasis. A long follow-up period aids in the assessment of bone lesions, because of the hypothesis that a metastatic lesion will change or increase in size over time, whereas a benign lesion will not. Follow-up periods in this study varied from no follow-up to 3 years.
The PET and bone studies should ideally be performed on the same day, or at least within several days of one another. The patient population selected for the study should reflect that seen in the general population where the two imaging tests would be utilized. A prospective study is preferable, because it can be designed to avoid many of the problems inherent in retrospective studies such as blinding and an objective and robust reference standard.

**Patient Selection**

We obtained approval of the institutional review board prior to the start of the study. The tumor registry and PET database at Duke University Medical Center (DUMC) were retrospectively reviewed to identify all patients with a pathologically proved new diagnosis of lung cancer, and staging with both whole-body PET and bone scan prior to the initiation of therapy. The review included 1,100 patients with lung cancer diagnosed between July 1998 and December 2000, and 2,000 consecutive patients who had a whole-body PET scan between September 2000 and August 2002. We chose these 2 data sets due to their ready availability. Though providing different starting points, both data sets should result in identification of all patients at DUMC who fulfill inclusion criteria for the specified time period.

This search identified 257 patients, 106 women and 151 men (age range 38-89 years; average age, 64.1 years) who fulfilled the entrance criteria and had sufficient followup to be considered eligible for analysis.
Interpretation of Imaging Studies

One hundred and twenty-eight patients received bone scan before PET scan, with an average delay of 11.2 days (range 0 – 133 days) between the scans. No patient had any therapeutic intervention between studies. While interpreting the scans, radiologists had unrestricted access to clinical information in the patients’ electronic medical record, including results or films of prior radiographic studies. This includes access to prior PET or bone scans. For the purposes of this study, we considered tumor in a rib adjacent to the primary tumor as direct extension and not true distant metastasis.

PET Imaging. Two hundred and fifty-six of the 257 patients underwent whole-body FDG-PET for staging of newly diagnosed lung cancer at DUMC. One patient received a PET study at an outside hospital and submitted the scans to DUMC for review.

PET imaging was performed on a GE Advance unit 2-D acquisition mode. Prior to January 2000, non-attenuation-corrected scans were initially obtained for 4 minutes per bed position from the level of the base of the brain through the middle of the thighs. A two bed position attenuation-corrected regional chest scan was then performed using 8 minutes for the emission scan and 10 minutes for the transmission scan at each bed position. The images were reconstructed using a Hann filter with a cutoff of 0.71/cm.

PET studies performed after January 2000 used emission images obtained for 4 minutes per bed position from the level of the base of the brain through the middle of the thighs. Transmission images using a Germanium–68 rod source
were then obtained for 2.5 minutes per bed position. The images were reconstructed using OSEM (2 iterations) and segmented attenuation correction. Reconstructed images were 128 x 128 x 35 with 4.30 mm x 4.30 mm x 4.25mm voxels. The transaxial, coronal, sagittal, and maximum intensity projection (MIP) were reviewed on the vendor-supplied workstation.

A single experienced nuclear medicine physician interpreted the PET study as positive for bone metastases if it contained a focal area of increased FDG uptake in bone greater than normal background bone activity. The nuclear medicine physician interpreted the PET studies as negative if there was no significant FDG uptake. The study investigators retrospectively reviewed each PET report and coded it as ‘positive’, ‘probable’, ‘unlikely’, or ‘negative’ for the presence of distant bone metastasis. A thoracic radiologist with over a decade of experience with FDG-PET in lung cancer re-examined any reports of PET that described indeterminate or equivocal results along with the original study. He then coded these studies into one of the aforementioned diagnostic categories for bone metastasis.

**Bone Scintigraphy.** Two hundred and thirty-four patients underwent radionuclide bone scintigraphy at DUMC. Twenty-three patients presented to DUMC with bone imaging performed at outside hospitals, the results of which we collected from clinic notes or reports by nuclear medicine physicians.

Whole-body images were obtained in the anterior and posterior projections starting 2-3 hours following the intravenous injection of 430 μCi/kg (15,910 kBq/kg; 30.0 mCi [1,110 MBq] maximum) of technetium 99m methylene...
diphosphonate. Either a BodyScan (Nuclear Medicine Group, Siemens Medical Group, Hoffman Estates, Ill) or a T-22 Twin Detector Nuclear Imaging System (SMV, Twinsburg, Ohio) was used for the radionuclide bone imaging studies.

Nuclear medicine physicians interpreted the bone scans according to standard clinical practice, using intensity, configuration, location, and number of foci of increased tracer activity. An experienced nuclear medicine physician interpreted the bone scintigraphy study as positive for bone metastasis if the radiotracer activity in the lesion was greater than normal bone. If there was no significant radiotracer uptake in the bones or radiotracer uptake was characteristic of benign disease, we coded the study as negative. If the location, distribution, and intensity of the radiotracer activity were indeterminate for differentiation between benign and metastatic disease, we coded the study as equivocal. We reviewed reports retrospectively and coded each as positive, probable, equivocal, unlikely, or negative for the presence of distant bone metastasis.33, 34

Clinical-pathologic Correlation

We analyzed all available clinical and pathologic information from patient records to determine the presence or absence of osseous metastatic disease at the time of initial staging. We entered all clinical data from clinic visits, radiographic studies, and followup into a database.

Confirmation of Bone Metastases. We considered patients to have bone metastases if any of the following criteria were present: positive bone biopsy (n = 6); progression of bony lesions seen on subsequent nuclear medicine studies.
(PET, n = 1; bone scan, n = 12); radiographic confirmation by other imaging modalities (n = 18); or ≥3 osseous lesions on initial nuclear medicine studies (PET scan, n = 8; bone scan, n = 3). Additional criteria included concordant initial PET and bone scans with similar distribution of osseous lesions (n = 4); or in the setting of discordant initial PET and bone scans without follow-up, confirmation of metastatic disease at non-osseous distant sites by PET or biopsy constituted sufficient evidence to support the presence of bone metastasis (n = 5).

**Absence of Bone Metastases.** We considered bone metastases to be absent if biopsy of a suspicious bone lesion was negative (n = 7) or suspicious bone lesions were not present on followup imaging studies within 1 month of initial staging (n = 112). In the absence of followup imaging studies, we considered patients with negative initial PET and bone scan and without clinical findings to have no bone metastasis (n = 81).

**Statistical Analysis**

We calculated sensitivity, specificity, positive predictive value, and negative predictive value along with 95% confidence intervals for PET and bone scintigraphy. We used the McNemar test was to compare the accuracy of PET and bone scintigraphy and calculated the weighted $\kappa$ statistic to determine their agreement. The $\kappa$ value is an adjustment of the percentage of agreement to accommodate for chance agreement. Clinicians have proposed the following values of $\kappa$ to qualitatively evaluate agreement: poor (<0), slight (0-0.20), fair (0.21-0.40), moderate (0.41-0.60), substantial (0.61-0.80), and excellent (0.80-
1.0). We performed all statistical analyses with the Stata version 7.0 (Stata Corporation, College Station, TX) software.

RESULTS

Patients

Lung cancer histology and stage distribution are listed in Table 1. One hundred and four (40%) of the 257 patients presented with stage IV disease, and 57 (22%) patients had osseous metastases by clinical-pathologic correlation.

For the calculation of sensitivity, specificity, and agreement, we combined all ‘positive’ and ‘probable’ PET and bone scan results into one category of positive. Likewise, we combined all ‘negative’ and ‘unlikely’ PET and bone scan results into one category of negative.

PET Scan

PET scan was positive in 61 (24%) and negative in 196 (76%) patients.

True positive PET. PET correctly identified bone metastases in 52 of the 57 (91%) patients with confirmed osseous disease. Bone scan found 21 patients with confirmed osseous Stage IV disease negative or equivocal for osseous metastasis, and PET correctly detected metastasis in 19 (90%) of these 21 patients. Bone was the only site of metastatic disease in 2 (10%) of these 19 cases. In 6 (32%) of the 19 patients, PET was the only initial staging study to identify distant metastasis and classify the patient as Stage IV.
**False positive PET.** In 9 patients, PET incorrectly identified 11 foci as positive for osseous metastasis. The lesions were noted in the spine \( (n = 5) \), ribs \( (n = 2) \), pelvis \( (n = 2) \), and acetabulum \( (n = 2) \). Bone scan accurately determined 7 of these 9 patients to be negative for osseous metastases. In no case did the false positive PET result affect patient management or the decision to perform surgery. All patients proceeded to receive surgical resection or neoadjuvant chemotherapy.

**False negative PET.** PET did not detect osseous metastases in 5 patients. Three of the 5 patients had a lesion outside of the region scanned by PET—one in the calvarium, one in the lateral femoral condyle, and one in the distal fibula. Bone was the only site of distant metastasis in 2 (40%) of the 5 patients, but in both cases initial staging CT scans also provided evidence for Stage IV disease. PET detected other non-osseous sites of distant metastasis in 2 (40%) of the 5 patients. Therefore, futile surgical intervention was not performed on any patient with a false negative PET scan.

The sensitivity, specificity, and positive and negative predictive values of PET were 91% (95% CI, 81% – 97%), 96% (95% CI, 92% - 98%), 85% (95% CI, 74% - 93%), and 97% (95% CI, 94% - 99%), respectively (Table 2).

**Bone Scan**

Bone scan was positive in 46 (18%), negative in 197 (77%), and equivocal in 14 (5%) patients.
**True positive Bone Scan.** Bone scintigraphy correctly identified 36 (63%) of the 57 patients with bone metastases as positive. Of the 21 patients with osseous metastases not considered positive by bone scintigraphy, 7 had equivocal scans. Among the 57 patients with Stage IV disease, PET missed metastatic lesions in 3 (5%) patients that bone scan found in the following locations: calvarium in 1 patient; mandible and fibula in 1 patient; and multiple lesions in the vertebral bodies, fibula, and femur in 1 patient. The lesion in the calvarium was outside the region scanned by PET, but routine staging head CT scan found it. Of the 3 patients, PET detected extraosseous, distant metastatic disease in 1 patient.

**False positive Bone Scan.** Bone scan incorrectly identified 15 benign foci as positive for metastatic disease in 10 patients. The locations of these foci were vertebral body \( (n = 3) \), skull \( (n = 3) \), rib \( (n = 3) \), sternum \( (n = 2) \), scapular tip \( (n = 1) \), costochondral junction \( (n = 1) \), mandible \( (n = 1) \), and vertebral pedicle \( (n = 1) \). PET correctly determined 9 of these 10 patients to be negative for osseous metastases. In one unfortunate case, the false positive bone scan resulted in the patient being considered ineligible for curative surgical resection.

**False negative Bone Scan.** Bone scan was negative in 14 patients with confirmed osseous metastases. Bone was the only site of distant metastasis in 2 of the 14 patients. PET was positive in all 14 patients and was the only initial staging study to provide evidence of distant metastatic disease in 6 cases.

**Equivocal Bone Scan.** Classification of all equivocal bone scan results as positive for metastatic disease yields the following values of sensitivity,
specificity, positive predictive value, and negative predictive value for bone scintigraphy: 75% (95% CI, 62% - 86%), 92% (95% CI, 87% - 95%), 72% (95% CI, 59% - 83%), and 93% (95% CI, 88% - 96%), respectively. Classification of equivocal bone scan results as negative yields corresponding values of 63% (95% CI, 49% - 76%), 95% (95% CI, 91% - 98%), 78% (95% CI, 64% - 89%), and 90% (95% CI, 85% - 94%) (Table 3).

**Summary of PET and Bone Scan Efficacy**

A summary of sensitivity, specificity, positive and negative predictive values of PET and bone scan is listed in Table 4. Application of PET and bone scan test characteristics to a hypothetical population of 1,000 patients is shown in Table 5. The rate of false-positive PET and bone scans is approximately 3-4%. False-negative PET scans account for 2% of results, whereas false-negative bone scans are a bit higher at 5.5%. Equivocal bone scans account for 5.5% of results.

**Likelihood Ratios**

Both PET and bone scan have likelihood ratios greater than 20 when ‘positive’ (Table 6). Only PET has a likelihood ratio less than 0.10 when ‘negative’. In the category of ‘probable’, the likelihood ratios of PET and bone scan are 10.5 and 4.7, respectively. In the category of ‘unlikely’, the likelihood ratios of PET and bone scan are 0.27 and 0.80, respectively. Ninety-three percent of PET results are in the ‘positive’ or ‘negative’ diagnostic category, whereas 74% of bone scan results are in the same categories.
Agreement and Accuracy

Agreement between PET and bone scintigraphy findings was calculated as \( \kappa = 0.510 \) (95% CI, 0.402 – 0.618) (Table 7).

Accuracy of PET (95%) is significantly greater than bone scintigraphy (90%) \((p < 0.05)\) when equivocal bone scan results are excluded from the analysis (Table 8). McNemar’s test is also significant when we consider all equivocal bone scintigraphy findings as either positive for metastatic disease \((p < 0.05)\) or negative for metastatic disease \((p < 0.001)\).

DISCUSSION

Once a new diagnosis of lung cancer has been established, an accurate assessment of the stage becomes crucial for providing prognosis information and deciding on treatment options. Physicians have traditionally determined clinical stage by a series of non-invasive radiologic and laboratory tests, which primarily demonstrate a gross anatomic distribution of disease. Conventional imaging studies are clearly less than optimal, as evidenced by the approximately 60% 5-year survival rates of patients initially diagnosed with early stage lung cancer.\(^8\) These patients typically undergo curative resection, because there is no evidence for metastases, but almost 40% later recur with disease, which was not detected by conventional imaging modalities.\(^8\) Thus the search continues for more accurate and efficient techniques to stratify patients for the most appropriate treatment protocols.
Over the last several years, FDG-PET has emerged as a novel and powerful imaging modality to complement CT scan for intrathoracic staging of lung cancer. PET has also received attention for improved pre-operative, extrathoracic staging in candidates for surgical resection. Several studies have demonstrated improved detection over conventional imaging of otherwise unsuspected distant metastases with PET in all sites except for brain.

The results of this study are in keeping with those of previous studies, but these results must be viewed skeptically because of sub-optimal study design. This study suggests that PET scan is more sensitive and as specific as bone scan for the evaluation of distant osseous metastasis in patients with newly diagnosed lung cancer. Calculation of a single sensitivity value for bone scan is not possible because of the presence of equivocal results, but the true sensitivity lies somewhere between the worst-case and best-case sensitivity values of 63% (95% CI, 49% - 76%) and 75% (95% CI, 62% - 86%), respectively. Sensitivity of PET (91%, 95% CI, 81% – 97%) is significantly greater than the worst-case estimate of bone scan sensitivity and trending toward statistical significance when compared to the best-case estimate. The agreement between PET and bone scan results with regard to distant osseous metastases is moderate.

The likelihood ratios associated with PET scan and bone scan favor the use of PET scan. Test results provided by bone scan are not as definitive as PET, as evidenced by the higher percentage of test results in the intermediate diagnostic categories of ‘probable’, ‘equivocal’, and ‘unlikely’. When ‘positive’, both PET and bone scan provide a large change in post-test probability with a likelihood
ratio greater than 20. But this value drops sharply to less than 5 for bone scan in the ‘probable’ diagnostic category. Likelihood ratios greater than 10 or less than 0.1 are considered the most useful in changing post-test probability. Likelihood ratios of 5-10 and 0.1-0.2 create moderate shifts in pre- to post-test probability. Likelihood ratios of 0.2-5 are generally of limited diagnostic utility. Bone scan had a likelihood ratio of 0.80 for ‘unlikely’ results, which comprised 15% of all bone scan results.

The results of this study are subject to a variety of biases. The retrospective selection of patients can easily result in selection or spectrum bias. Because DUMC is a major referral center, the expectation is that patients diagnosed with lung cancer by outside physicians will receive a PET scan at Duke only if no evidence of simultaneous metastatic disease is visualized by outside bone scan, thoracic CT, or brain CT scan. Such selection would reduce the prevalence of bone metastases in the study population and thus increase the difference in diagnostic efficiency of the two studies. An artificially low prevalence of metastatic disease in the study population spectrum would create a corresponding increase in the diagnostic efficiency of the two tests. The clinical reality may be that no statistically significant difference between the two tests exists among patients in the general population. However, the study population had a clinical stage distribution and prevalence of osseous metastases similar to lung cancer patients in the general population, of whom only 15% have localized disease and 20% osseous metastases at presentation. Therefore, it is unclear whether selection bias played a large role in influencing the results of this study.
It is possible that patient selection somehow favored PET scan and underestimated the accuracy of bone scintigraphy by an unknown mechanism related to selection bias.

Verification of the presence or absence of osseous metastasis at the time of initial staging with PET and bone scan is the most obvious source of bias. If this study has a fatal flaw, it is the lack of a rigorous and uniformly-applied diagnostic gold standard. While bone biopsy of the most suspicious lesions is the gold standard, physicians use it sparingly because of patient discomfort, and in many cases, results of the test would not change clinical management. Patients with initially positive PET or bone scan findings are more likely to receive subsequent investigation of bone lesions than those with negative scans, creating another form of verification bias known as work-up bias.34

Use of the results from initial PET and bone scans to verify metastatic disease creates incorporation bias. However, in this and all previous studies, the researchers employed the results of the PET and bone scans to varying degrees as evidence to determine the presence of metastasis or benign bone lesions. Incorporation bias refers to this use of circular logic. Other studies used a long follow-up period and the results of multiple bone scans or PET scans to minimize the incorporation bias of a single test result, preferring to rely on several results.

Incorporation bias and verification bias are difficult problems to rectify, since patients find the gold standard undesirable, and the existing diagnostic tests of PET and bone scan seem to have fairly good sensitivity and specificity. Given the difficulty of relying on the gold standard of biopsy, an alternate gold standard
using a combination of long-term follow-up and other imaging modalities such as plain film and MR may be feasible.

For patients diagnosed initially with no metastatic disease, it is difficult to prove the absence of such disease. Even the absence of bone lesions in follow-up radiographic studies does not conclusively prove the absence of osseous metastases at the time of initial staging, because these patients may simply be responding successfully to systemic chemotherapy or local radiation.

Agreement between the results of PET scan and bone scintigraphy studies was moderate. This result is also subject to bias, namely test review bias. Because unblinded radiologists reading the initial PET scan and bone scan may have used the result of the first study to interpret the second, the calculated $\kappa$ statistic may be artificially elevated. It is unclear how selection bias in this study affects the $\kappa$ statistic for agreement of PET and bone scan findings.

Of note is the effect of false-positive and false-negative results of PET scan and bone scan on patient management. False negative PET results had no adverse effect on patients, because all 5 (2%) patients had either simultaneous, non-osseous, extrathoracic metastases or the initial staging CT scans found the osseous metastases missed by PET. Similarly, false positive PET results did not result in any of the 7 (3%) patients being denied curative surgery. All patients proceeded to receive curative resection or neoadjuvant chemotherapy. However, in the case of false positive bone scans, 1 of the 10 (4%) patients was immediately ruled out for curative surgery by the thoracic surgeon. Moreover, in 6 of the 14
(5%) patients with a false negative bone scan, PET scan was the only staging study to provide evidence of Stage IV disease.

Cost-effectiveness analyses have demonstrated advantages to the addition of PET in the work-up of solitary pulmonary nodules.\textsuperscript{44,45} Lung cancer is the only neoplasm for which Medicare has approved reimbursement for FDG-PET in diagnosis, staging, and re-staging. If fused PET-CT does indeed become the gold standard of the future in lung cancer staging evaluations, then there is a potential cost savings of eliminating bone scan and relying completely on fused PET-CT scans. However, to date there is no conclusive evidence that PET can replace bone scintigraphy. A future study of better quality may eventually change the role of bone scan to that of a second-line, alternative imaging modality used only in cases of diagnostic uncertainty. Further research of this area should be performed in a well-designed, prospective, clinical trial to minimize bias.

A randomized, controlled trial can be used to evaluate the effect of PET scan and bone scan in the management of bone metastases and overall mortality of lung cancer patients. The three arms of the trial would be PET scan only, bone scan only, and both PET and bone scan for the detection of osseous metastasis. In an alternative, non-randomized trial design, investigators can perform both PET scan and bone scan on all patients. Investigators would selectively blind clinicians to the results of one or the other test and ask the clinicians to record their clinical management plans. Then, the clinician would record clinical management plans based on the information from both tests.\textsuperscript{46} Investigators
would determine which strategy resulted in the highest sensitivity and specificity. This type of design can be used in a cost-effective analysis.

Blinding is imperative in studies of diagnostic efficiency.\textsuperscript{36,37} For this particular clinical question of assessing bone metastasis in lung cancer patients, the most important aspect of the study design is an \textit{a priori} protocol for the work-up and verification of suspicious osseous foci as benign or metastatic. When bone biopsy is not performed, rigorous diagnostic criteria should be employed as a substitute gold standard. The substitute gold standard will most likely require the judicious use of other imaging modalities and an extended follow-up period of approximately one year or more. Special attention is required to determine the optimal period of follow-up with PET scan and bone scan, the effect of localized radiation therapy or systemic chemotherapy on the subsequent detection of osseous metastases, and the clinical significance of equivocal bone scan results.
## TABLE 1
Histological Diagnosis and Stage Distribution of Study Population

<table>
<thead>
<tr>
<th>Histology</th>
<th>No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-small cell lung cancer, otherwise unspecified</td>
<td>91</td>
</tr>
<tr>
<td>Adenocarcinoma</td>
<td>72</td>
</tr>
<tr>
<td>Squamous cell carcinoma</td>
<td>59</td>
</tr>
<tr>
<td>Small cell lung carcinoma</td>
<td>18</td>
</tr>
<tr>
<td>Large cell carcinoma</td>
<td>10</td>
</tr>
<tr>
<td>Bronchioalveolar carcinoma</td>
<td>3</td>
</tr>
<tr>
<td>Adenosquamous carcinoma</td>
<td>1</td>
</tr>
<tr>
<td>Carcinosarcoma</td>
<td>1</td>
</tr>
<tr>
<td>Carcinoid</td>
<td>1</td>
</tr>
<tr>
<td>Pleiomorphic carcinoma</td>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Stage Distribution</th>
<th>No. (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stage I</td>
<td>33 (13)</td>
</tr>
<tr>
<td>Stage II</td>
<td>24 (9)</td>
</tr>
<tr>
<td>Stage IIIA</td>
<td>40 (16)</td>
</tr>
<tr>
<td>Stage IIIB</td>
<td>56 (22)</td>
</tr>
<tr>
<td>Stage IV</td>
<td>104 (40)</td>
</tr>
<tr>
<td>Clinical Pathologic Findings</td>
<td>PET Findings</td>
</tr>
<tr>
<td>------------------------------------------</td>
<td>--------------</td>
</tr>
<tr>
<td>Confirmed bone metastases</td>
<td>Positive</td>
</tr>
<tr>
<td>Absence of bone metastases</td>
<td>Negative</td>
</tr>
</tbody>
</table>

Sensitivity

52/57 = 91% (95% CI, 81% - 97%)

Specificity

191/200 = 96% (95% CI, 92% - 98%)

Positive Predictive Value

52/61 = 85% (95% CI, 74% - 93%)

Negative Predictive Value

191/196 = 97% (95% CI, 94% - 99%)
### TABLE 3
Radionuclide Bone Scintigraphy versus Clinical Pathologic Findings

<table>
<thead>
<tr>
<th>Clinical Pathologic Findings</th>
<th>Bone Scintigraphy Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Confirmed bone metastases</td>
<td>Positive</td>
</tr>
<tr>
<td>36</td>
<td>14</td>
</tr>
<tr>
<td>Absence of bone metastases</td>
<td>10</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Clinical Pathologic Findings</th>
<th>Bone Scintigraphy Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Confirmed bone metastases</td>
<td>Positive</td>
</tr>
<tr>
<td>43</td>
<td>14</td>
</tr>
<tr>
<td>Absence of bone metastases</td>
<td>17</td>
</tr>
</tbody>
</table>

*Equivocal bone findings are considered positive for metastatic disease*

- **Sensitivity**
  - Confirmed bone metastases: $\frac{43}{57} = 75\%$ (95% CI, 62% - 86%)
  - Absence of bone metastases: $\frac{183}{200} = 92\%$ (95% CI, 87% - 95%)

- **Specificity**
  - Confirmed bone metastases: $\frac{190}{200} = 95\%$ (95% CI, 91% - 98%)
  - Absence of bone metastases: $\frac{190}{211} = 90\%$ (95% CI, 85% - 94%)

- **Positive Predictive Value**
  - Confirmed bone metastases: $\frac{43}{60} = 72\%$ (95% CI, 59% - 83%)
  - Absence of bone metastases: $\frac{183}{197} = 93\%$ (95% CI, 88% - 96%)

- **Negative Predictive Value**
  - Confirmed bone metastases: $\frac{36}{46} = 78\%$ (95% CI, 64% - 89%)
  - Absence of bone metastases: $\frac{190}{200} = 95\%$ (95% CI, 91% - 98%)
<table>
<thead>
<tr>
<th></th>
<th>PET</th>
<th>Bone scan (equivocal→positive)</th>
<th>Bone scan (equivocal→negative)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensitivity</td>
<td>0.91 (0.81,0.97)</td>
<td>0.75 (0.62,0.86)</td>
<td>0.63 (0.49,0.76)</td>
</tr>
<tr>
<td>Specificity</td>
<td>0.96 (0.92,0.98)</td>
<td>0.92 (0.87,0.95)</td>
<td>0.95 (0.91,0.98)</td>
</tr>
<tr>
<td>Positive predictive value</td>
<td>0.85 (0.74,0.93)</td>
<td>0.72 (0.59,0.83)</td>
<td>0.78 (0.64,0.89)</td>
</tr>
<tr>
<td>Negative predictive value</td>
<td>0.97 (0.94,0.99)</td>
<td>0.93 (0.88,0.96)</td>
<td>0.90 (0.85,0.94)</td>
</tr>
</tbody>
</table>

95% confidence intervals listed in parantheses
### TABLE 5
Results of PET and Bone scan in Hypothetical Population (n = 1000)

<table>
<thead>
<tr>
<th>Clinical Pathologic Findings</th>
<th>PET Findings</th>
<th>Bone Scintigraphy Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Positive</td>
<td>Negative</td>
</tr>
<tr>
<td>Confirmed bone metastases</td>
<td>202</td>
<td>20</td>
</tr>
<tr>
<td>Absence of bone metastases</td>
<td>35</td>
<td>743</td>
</tr>
<tr>
<td>Confirmed bone metastases</td>
<td>140</td>
<td>55</td>
</tr>
<tr>
<td>Absence of bone metastases</td>
<td>39</td>
<td>712</td>
</tr>
</tbody>
</table>
TABLE 6
Comparison of PET and Bone scan Likelihood Ratios

<table>
<thead>
<tr>
<th>Clinical Pathologic Findings</th>
<th>PET Findings</th>
<th>Bone Scintigraphy Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Confirmed bone metastases</td>
<td>Positive</td>
<td>Probable</td>
</tr>
<tr>
<td>Absence of bone metastases</td>
<td>49</td>
<td>3</td>
</tr>
</tbody>
</table>

\[
LR_{\text{positive}} = 21.5 \\
LR_{\text{probable}} = 10.5 \\
LR_{\text{equivocal}} = 0.27 \\
LR_{\text{negative}} = 0.08 \\
\]

\[
LR_{\text{positive}} = 24.6 \\
LR_{\text{probable}} = 4.7 \\
LR_{\text{equivocal}} = 3.5 \\
LR_{\text{unlikely}} = 0.80 \\
LR_{\text{negative}} = 0.16 \\
\]

22% 2% 5% 71%
**TABLE 7**

Agreement Between PET and Radionuclide Bone Scintigraphy Findings

<table>
<thead>
<tr>
<th>PET Findings</th>
<th>Bone Scintigraphy Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Positive</td>
</tr>
<tr>
<td>Positive</td>
<td>34</td>
</tr>
<tr>
<td>Negative</td>
<td>12</td>
</tr>
</tbody>
</table>

weighted $\varnothing$ statistic = 0.510 (95% CI, 0.402 – 0.618)
### TABLE 8
Accuracy of PET and Radionuclide Bone Scintigraphy Findings

<table>
<thead>
<tr>
<th>PET Findings</th>
<th>Bone Scintigraphy Findings</th>
<th>Correct</th>
<th>Incorrect</th>
<th>Equivocal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correct</td>
<td>Correct</td>
<td>209</td>
<td>23</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>Incorrect</td>
<td>10</td>
<td>1</td>
<td>3</td>
</tr>
</tbody>
</table>

McNemar’s test (ignoring equivocal results)
McNemar Chi Square  $p = 0.035$

<table>
<thead>
<tr>
<th>PET Findings</th>
<th>Bone Scintigraphy Findings</th>
<th>Correct</th>
<th>Incorrect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correct</td>
<td>Correct</td>
<td>214</td>
<td>29</td>
</tr>
<tr>
<td></td>
<td>Incorrect</td>
<td>12</td>
<td>2</td>
</tr>
</tbody>
</table>

*equivocal bone findings are considered positive for metastatic disease

McNemar’s test (counting all equivocal results as positive for metastatic disease)
McNemar Chi Square  $p = 0.012$

<table>
<thead>
<tr>
<th>PET Findings</th>
<th>Bone Scintigraphy Findings</th>
<th>Correct</th>
<th>Incorrect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correct</td>
<td>Correct</td>
<td>215</td>
<td>28</td>
</tr>
<tr>
<td></td>
<td>Incorrect</td>
<td>11</td>
<td>3</td>
</tr>
</tbody>
</table>

*equivocal bone findings are considered negative for metastatic disease

McNemar’s test (counting all equivocal results as negative for metastatic disease)
McNemar Chi Square  $p < 0.001$
References


