Annals of Nuclear Medicine Vol. 18, No. 6, 483–488, 2004

A problem in diagnosing N3 disease using FDG-PET in patients with lung cancer —High false positive rate with visual assessment—

Masaki HARA,* Norio SHIRAKI,* Masato ITOH,* Yuta SHIBAMOTO,* Akihiko IIDA,**
Masami NISHIO*** and Tsuneo TAMAKI***

*Department of Radiology, Nagoya City University Graduate School of Medical Sciences
**Department of Radiology, Nagoya City Rehabilitation Center
***Department of Radiology, Nagoya Kyoritu Hospital

Objective: To evaluate the accuracy of diagnosing N3 disease using positron emission tomography (PET) with 2-[fluorine-18]fluoro-2-deoxy-D-glucose (FDG) in patients with pulmonary disease. ***Subjects and Methods:*** Twenty patients diagnosed as FDG-PET N3 were enrolled. On FDG-PET, lymph nodes were considered to be positive when increased uptake as compared with that of the surrounding mediastinum was visually observed, or the mean standardized uptake ratio (SUR) was more than 2, 2.5, or 3. On CT, lymph nodes exceeding 1 cm in the shortest diameter were regarded as positive. ***Results:*** The PET result was true positive (TP) in 2 patients and false positive (FP) in 18 with an overall accuracy (OA) of 10% using visual criteria. Using an SUR of more than 2.5, the result was TP in 2, FP in 3, and true negative (TN) in 15, the false negative (FN) in 0, with an OA of 85%. CT diagnosis was TP in 2, FP in 9, and TN in 9 with an OA of 55%. The accuracy using the SUR criteria of more than 2.5 was superior to that of CT. ***Conclusion:*** Of 20 patients with the diagnosis of PET N3, we found frequent over-diagnosis in nodal staging using the visual criteria.

Key words: FDG-PET, nodal staging, false positive, N3, lung cancer

INTRODUCTION

POSITRON EMISSION TOMOGRAPHY (PET) with 2-[fluorine-18]fluoro-2-deoxy-D-glucose (FDG) has been accepted as a useful imaging modality that is completely noninvasive and available for obtaining information on the nodal status and distant metastasis.[1–7] Although FDG-PET is highly sensitive to detect malignant foci, particularly in patients with poorly-differentiated tumor,[8] increased accumulation to benign lesions such as inflammation, infection, and granulation has produced some serious problems in evaluating lung cancer staging.[9,10]

Until the 1950's, tuberculosis was quite common all over Japan and the disease rate was as high as over 500 per 100,000 people.[11] At the present time, many aged patients with lung cancer in our country have a past history of pulmonary tuberculosis. Thus, we suspected that these old granulomatous diseases might affect the accuracy of FDG-PET for diagnosing nodal status, due to high accumulation of FDG in benign as well as malignant lesions. In this study we report a limitation of FDG-PET for nodal staging in lung cancer patients living in a region with a high incidence of a past tuberculosis.

Received February 25, 2004, revision accepted April 21, 2004.

For reprint contact: Masaki Hara, M.D., Department of Radiology, Nagoya City University Graduate School of Medical Sciences, 1 Kawasumi, Mizuho-cho, Mizuho-ku, Nagoya 467–8601, JAPAN.

SUBJECTS AND METHODS

Patients

Of all 60 patients with thoracic diseases who underwent FDG-PET at our institution between November 1998 and December 2000, 20 patients (13 males and 7 females; aged 53–79 years; mean age, 68 years) were diagnosed to have N3 disease and were enrolled in this study. We reviewed imaging and clinical findings in all of these patients retrospectively. Although all patients were initially considered to have lung cancer, 7 proved to have

benign disease later; 2 each had inflammation, granuloma, and tuberculosis, and one had sarcoidosis. The other 13 patients had a primary lung cancer. Of them, 8 had adenocarcinoma, 3 had squamous cell carcinoma, and one each had large cell carcinoma and simultaneous adenocarcinoma and squamous cell carcinoma. Later, 2 patients proved to have N3 disease, whereas eleven proved to have N0 disease. All 2 N3 and 2 N0 patients had confirmation of their nodal status by careful follow-up using CT and the other nine N0 patients were diagnosed both clinically using CT and pathologically using specimens of lymph nodes resection except for contralateral hilar nodes. The follow-up period of the N0 patients ranged from 30 to 52 months, mean 40 months.

FDG-PET

FDG-PET was performed using a PCT3600W unit (Hitachi Medical Co., Kashiwa, Japan). The PCT3600W unit produced 6.0-mm-thick image planes (8 direct planes and 7 cross planes). The resolution of the scanner at full width at half maximum was 6.5 mm and longitudinal field view was 10.5 cm. Image processing and reconstruction were performed with a computer system (HARP, Hitachi Advanced Radionuclide Processor, Kashiwa, Japan).

All patients fasted for at least 6 hours before imaging. Transmission scan at 20 minutes was obtained before emission scan by using a rotating germanium 68 pin source. One transverse emission scan at 30 minutes including hilar and mediastinal region was then obtained starting 50 minutes after administration of 300 MBq FDG.

FDG accumulation within the regional lymph nodes on attenuated corrected images was evaluated independently on hard copy images and reviewed by 3 experienced radiologists (M.H., A.I., N.S.), and consensus scores were recorded. Lymph nodes were considered to be positive when 1) increased uptake as compared with that of the surrounding mediastinum was visually observed, or 2) the mean standardized uptake ratio (SUR) was more than 2, 2.5, or 3. The nodes were judged to be negative if 1) the activity was equal to or less than that of the adjacent mediastinum, or 2) the mean SUR was 2, 2.5, and 3 or less. A circular region of interest was placed on the emission image with maximum FDG uptake. Nodal status was classified as N1 if the positive node was located in the ipsilateral hilum, N2 if it was in the ipsilateral mediastinum, and N3 if it was in the contralateral hilum or mediastinum. SUR was calculated as follows:

$$\text{SUR} = \frac{\text{mean ROI uptake (MBq/m}l\text{)}}{\text{injected dose (MBq)/weight (kg)}}$$

CT

Thoracic CT examinations were performed using a HiSpeed Advantage SG scanner (GE Medical Systems, Milwaukee, USA) or Somatom Plus 4 (Siemens Medical Instrumentation, Erlangen, Germany). All 20 patients underwent enhanced helical CT 50 seconds after intravenous administration of 100 m*l* of non-ionized contrast material at a rate of 1.2–1.5 m*l*/sec with 2–7 mm collimation (most commonly 5 mm) at pitch 1, from the apex to just below the inferior pulmonary vein. On CT, lymph nodes exceeding 1 cm in the shortest diameter were regarded as positive for metastasis. All CT images were reviewed until a consensus was reached by 3 experienced radiologists (M.H., N.S., H.O.), and the nodal stage was diagnosed.

Nodal Stage and Radiological-Histopathological or -Clinical Correlation

The nodal status was confirmed by nodal dissection during thoracotomy (n = 8 [40%]), or determined from the clinical course (n = 12 [60%]) in all 7 patients with benign diseases and 5 with lung cancers (3 patients with marked lymph node enlargement and 2 with no marked interval change at least for 2 years on CT). Ipsilateral hilar and mediastinal dissection was performed during thoracotomy.

The accuracy of PET staging of the N3 disease using the visual and the SUR criteria was compared with that of enhanced CT. In addition, the association between the false-positivity rate of PET in diagnosing N3 and the size of the primary lesion, smoking history, or past history of pulmonary disease was evaluated.

RESULTS

CT Stage

In all 20 patients, CT finding was true positive (TP) in 2 patients, false positive (FP) in 9 patients, and true negative (TN) in 9 patients with an overall accuracy (OA) of 55%, sensitivity of 100%, specificity of 50%, positive predictive value (PPV) of 18%, and negative predictive value (NPV) of 100% (Table 1). In 13 patients with cancer, CT result was TP in 2, FP in 5, and TN in 6 and accuracy was elevated with an OA of 62%, sensitivity of 100%, specificity of 55%, PPV of 29%, and NPV of 100% (Table 2).

FDG-PET Stage

Using visual criteria in the 20 patients, PET result was FP in 18 and TP in only 2, yielding a low OA and PPV of 10%. In 13 patients with malignant disease, PET result was TP in 2 patients and FP in 11, with an OA and PPV of 15% (Tables 1, 2). Both results were far inferior to those of contrast-enhanced helical CT.

Among benign nodes in this study, calculated mean SURs for contralateral lymph nodes were distributed from 1.6 to 3.1 (mean; 2.3). In two metastasized cases, mean SURs of contralateral lymph nodes were 2.9 and 3.0.

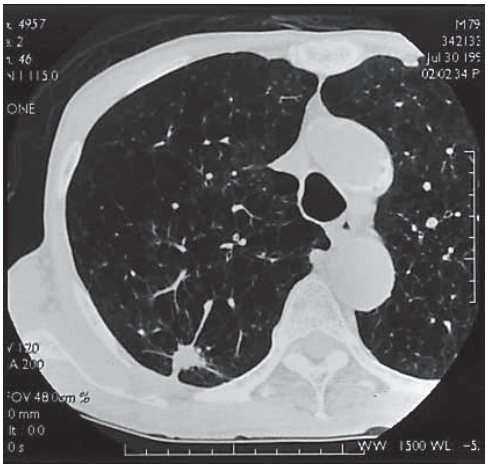
Using an SUR cut-off value of 2, the diagnosis was TP in 2 patients, FP in 10, and TN in 8, with an OA of 50%, sensitivity of 100%, specificity of 44%, PPV of 17%, and

Table 1 Accuracy of nodal staging (n = 20)

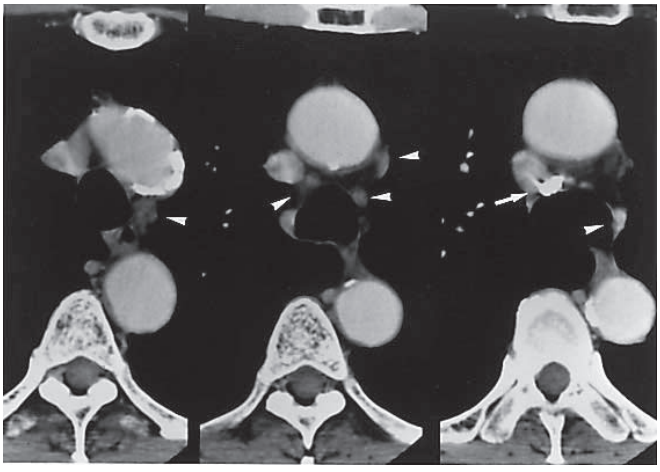
	PET (visual)	PET (SUR > 2)	PET (SUR > 2.5)	CT
True Positive	2	2	2	2
True Negative	0	8	15	9
False Positive	18 (85%)	10 (50%)	3 (15%)	9 (45%)
False Negative	0	0	0	0
Positive Predictive Value	2/20 (10%)	2/12 (17%)	2/5 (40%)	2/11 (18%)
Overall Accuracy	2/20 (10%)	10/20 (50%)	17/20 (85%)	11/20 (55%)

Table 2 Accuracy of nodal staging in patients with lung cancer (n = 13)

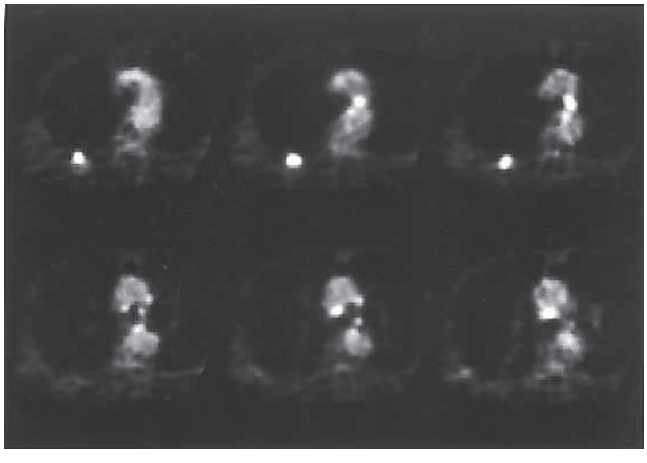
	PET (visual)	PET (SUR > 2)	PET (SUR > 2.5)	CT
True Positive	2	2	2	2
True Negative	0	5	10	6
False Positive	11 (77%)	6 (46%)	1 (38%)	5 (38%)
False Negative	0	0	0	0
Positive Predictive Value	2/13 (15%)	2/8 (25%)	2/3 (67%)	2/7 (29%)
Overall Accuracy	2/13 (15%)	7/13 (54%)	12/13 (92%)	8/13 (62%)



A



B



C

Fig. 1 A 79-year-old man with well-differentiated adenocarcinoma in the right upper lobe. A: Transverse CT image shows a 1.5 × 1.0 cm nodule with spiculations on a background of severe emphysema. B: Enhanced CT at the level of the carina shows numerous tiny mediastinal nodes less than 1 cm in diameter with (*arrows*) or without (*arrowheads*) calcification. C. Transverse PET scans through the same region as in A show areas of higher FDG uptake than the background mediastinum level both in the upper lobe nodule and lymph nodes regardless of calcification with SURs of 1.9 and 1.2–1.7, respectively. A partial resection and sampling of the subcarinal nodes revealed well-differentiated adenocarcinoma in the nodule with no lymph nodes metastases histopathologically. There has been no recurrence and no lymphadenopathy on CT performed 30 months later. This represents a false-positive study with the visual criteria but a true-negative study using the SUR criteria.

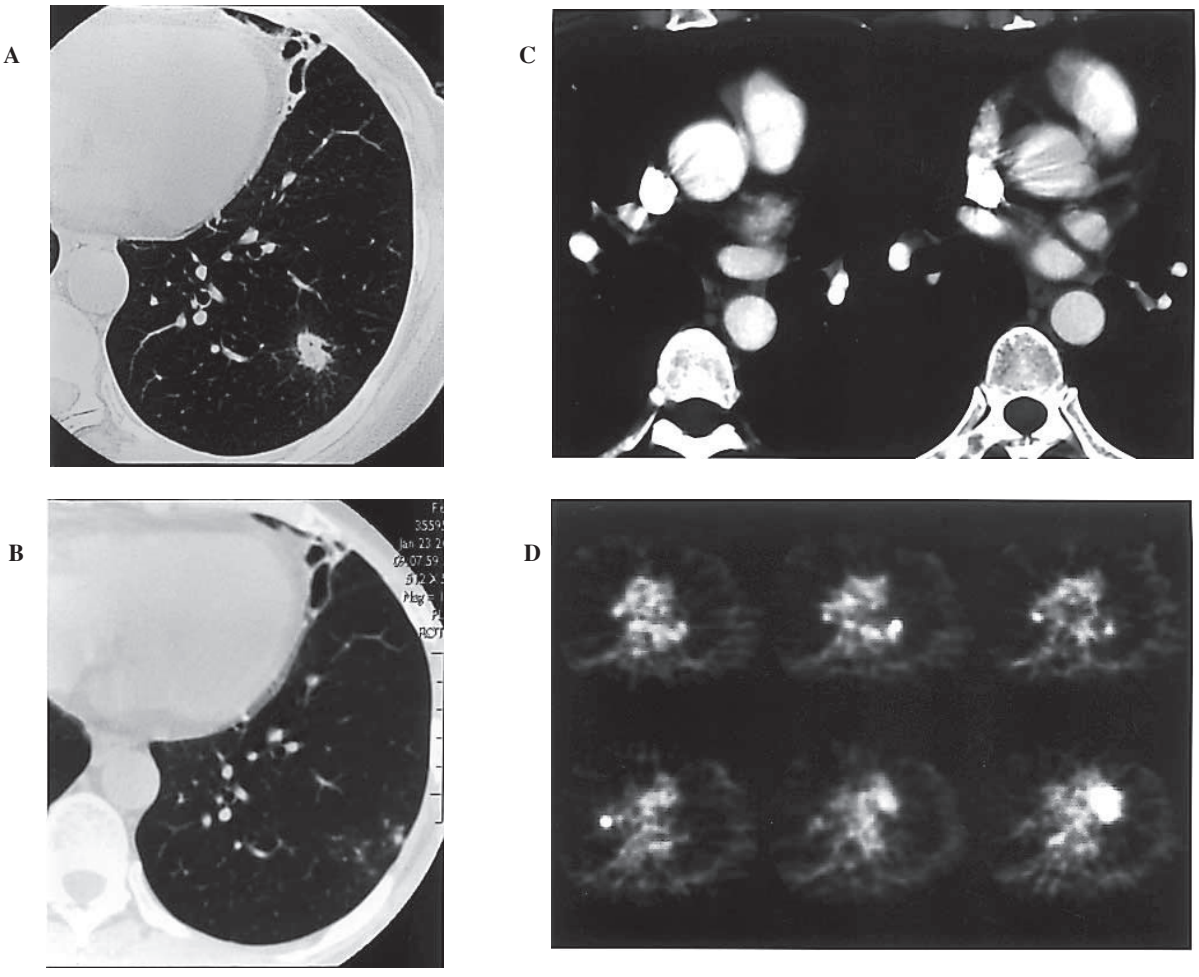


Fig. 2 A 64-year-old woman with a left lower lobe nodule that was incidentally detected on low-dose screening CT for lung cancer. A: Transverse CT image shows a 1.3 × 1.0 cm nodule with air bronchogram and irregular margin. B: The nodule had regressed spontaneously on CT 5 months later. C: CT of the hilum shows only small nodes less than 1 cm in short axis. D: Transverse PET scans at the level of the hilum show areas of increased FDG uptake compared to the background level with the SURs between 2.0 and 2.7. This represents a false-positive study with both the visual and SUR criteria.

NPV of 100%. In 13 patients with malignancy, PET result was TP in 2, FP in 6, and TN in 5 with an OA of 54%, sensitivity of 100%, specificity of 45%, PPV of 25%, and NPV of 100%, similar to those of contrast-enhanced helical CT (Tables 1, 2).

Using an SUR cut-off value of 2.5, the diagnosis was TP in 2 patients, FP in 3, TN in 15, and FN in 0, with an OA of 85%, sensitivity of 100%, specificity of 83%, PPV of 40%, and NPV of 100%. In 13 patients with malignancy, PET result was TP in 2, FP in 1, and TN in 10 with an OA of 92%, sensitivity of 100%, specificity of 91%, PPV of 67%, and NPV of 100%, superior to those of contrast-enhanced helical CT (Tables 1, 2).

Using an SUR cut-off value of 3.0, the diagnosis was TP in 0 patients, FP in 1, TN in 17, and FN in 2, with an OA of 85%, sensitivity of 0%, specificity of 94%, PPV of 0%, and NPV of 94% with increasing FN result. In 13 patients with malignancy, PET result was TP in 0, FP in 1, TN in 10, and FN in 2, with an OA of 77%, sensitivity of 0%, specificity of 91%, PPV of 0%, and NPV of 83%. On this setting all metastasized nodes resulted in FN.

In 18 patients without lymph node metastases, the mean SUR for FDG was 2.2 (range 1.2–3.1) with a smoking history and 2.0 (1.4–2.7) with no smoking history, showing no marked difference ($p = 0.34$, Mann-Whitney U test). In 11 patients with malignant disease, the mean SUR for FDG was 2.2 (range 1.2–3.1) with a smoking history and 1.7 (1.4–2.0) with no smoking history, a showing slightly high SUR scores in the former, but without statistical significance ($p = 0.083$). The mean SUR in patients with and without a history of pulmonary disease was 2.3 (1.7–2.9) in 13 lymph nodes and 2.1 (1.2–3.1) in 18 lymph nodes, respectively, showing no significant difference ($p = 0.22$). There was no significant difference either in the SUR between patients with lymph node calcification (mean SUR 1.9, 1.7–2.2, n = 2) and those

without (mean 2.2, 1.2–3.1, n = 29) (p = 0.52).

DISCUSSION

Accurate diagnostic staging is essential for choosing an appropriate treatment for primary lung cancers. However, diagnostic accuracy for hilar and mediastinal nodal status has been problematic. The limitation of CT and MRI is owing to its morphological basis of diagnosis; there exist large reactive benign nodes as well as small metastasized nodes. Preliminary evaluation using MR short inversion time inversion recovery (STIR) sequence has also been reported to show a high positive predictive value of 81%. However, it has not yet been performed in general practice.[12] For these reasons, diagnostic mediastinoscopy is still an indispensable tool to determine whether the patient is a candidate for surgical resection or not. PET with FDG has been considered to be a noninvasive technique that provides metabolic information appropriate for evaluating malignant lesions.[1–7] Initially several excellent results were reported,[1–7] but further investigation revealed a problem, namely, a considerably high rate of false positivity.[13–15] In this study we focused on the accuracy of PET N3 diagnosis, which essentially affects the clinical management of the patients.

The frequency of false positivity among PET N3 patients reached 85% (11/13) in patients with malignant disease using visual criteria. This accuracy is much lower than that of previous reports,[1–7] but is similar to that of a recent report from eastern Asia in patients with esophageal cancer that showed a very low PPV of 3% (1/33) in diagnosing hilar nodes metastasis.[15] We speculated regarding some causes of such false positive result. First, physiological, particularly, inhomogeneous FDG uptake to the mediastinum resulted in the over-diagnosis of nodal status.[9] Second, reactive lymphadenopathy related to peripheral malignancy such as sarcoid reaction and secondary obstructive pneumonia should cause an increased accumulation to hilar and mediastinal lymph nodes.[9,10] And the final and perhaps most important reason could be presence of organized lymphadenopathy due to old infection and/or the formerly widespread granulomatous disease, such as tuberculosis, in the Japanese population.[11] It seems probable that there are abundant macrophages even in reactive or organized lymph nodes.[10,15,16] The excellent sensitivity to glucose metabolism of FDG-PET may worsen the accuracy of diagnosis for nodal status. The report that even asymptomatic inflammation or smoking could produce an increased accumulation of FDG to the hilar and mediastinal lymph nodes would support our results.[16,17]

The criteria for diagnosing nodal status have been controversial. Visual criteria have been a usual method[1–3,5–7] and semi-quantitative value such as standardized uptake value (SUV) or SUR was reported not to be superior to a visual one.[13,17] Lymph nodes may be too small to place ROIs and calculate reliable values.[1] In contrast, semi-quantitative analysis was reported to be useful in reducing false-positive cases improving overall accuracy. An appropriate setting of diagnostic criteria using the maximum semi-quantitative analysis (a threshold SUV of 4.2) has been reported to be useful to improve overall accuracy,[3] but a direct comparison with visual criteria was not performed in that study. Mean SURs have been reported to be more accurate than maximum ones because of a relative increase in the contribution of noise in regions of low average counts.[18] Since in our series all contralateral lymph nodes showed a less than maximum semi-quantitative value of 4.2, we performed mean semi-quantitative analysis step-by-step from 2.0. The accuracy, in our series, improved using an SUR cut-off value of 2.5 as the criteria of malignancy. However, because the SURs distributed within a narrow range with considerable overlap and the method of ROI calculation for small targets essentially contains an operator-dependent bias to some extent, we think that the criteria of SUR would have a difficulty to differentiate benign from malignant nodes in individual clinical cases. Although an SUR cut-off value of 2.5 should decrease the number of false-positive cases, it may result in worsening the negative predictive value that has been one of the most outstanding advantages of FDG-PET.

This result also suggests that in the lesions with a low SUR such as well-differentiated adenocarcinoma FDG-PET may carry a risk of underestimation as well as a risk of over staging by increased accumulation in benign lesions. The limitation of our study is the relatively high frequency of clinical diagnosis for non-metastatic nodes, performed in 2 (15%) of 13 patients with malignant disease. Indolent metastasis might have been concealed among the false-positive group and might have reduced the accuracy in spite of the adequate follow-up period.

We agree that negative FDG-PET of the bilateral hilum and mediastinum is the most reliable criteria for the diagnosis of clinical N0 status except for less frequent intranodal micrometastases.[14,19] However, we suggest that FDG N0 may be infrequent in the area permeated by widespread granulomatous diseases such as tuberculosis in the past and/or present, particularly, in patients with thoracic diseases. Even though our series was too small to make any conclusion, we believe that symmetrical and bihilar accumulation higher than the background level in the mediastinum should not be considered as positive using visual criteria particularly in patients with a relatively low FDG accumulation in the primary lesion. Experienced nuclear medicine physicians may clinically diagnose like this way, but this point has not been mentioned in the past literature.

In this study we could not find any significant difference in the SUR between the benign and malignant nodes with regard to smoking history, past history of pulmonary diseases or degree of calcification within nodes. Although

FDG accumulation in some intratumoral components such as macrophages, granulation tissue and inflammatory cells has been reported,[20] larger and prospective studies on radiological-pathological correlations are needed to clarify the etiology of increased FDG uptake by non-metastatic regional lymph nodes.

In summary, of 20 patients with the diagnosis of PET N3, we found frequent over-diagnosis in nodal staging using the visual criteria. In our series an SUR cut-off value of 2.5 as the criteria of malignancy improved the accuracy for nodal staging of N3, but it may result in reducing the negative predictive value. We suggest that symmetrical hilar and mediastinal accumulation only slightly higher than the background mediastinum level should not be considered to be positive, particularly with a background of formerly widespread granulomatous disease.

REFERENCES

1. Wahl RL, Quint LE, Greenough RL, Meyer CR, White RJ, Orringer MB. Staging of mediastinal non-small cell lung cancer with FDG PET, CT, and fusion image: preliminary prospective evaluation. *Radiology* 1994; 191: 371–377.
2. Patz EF Jr, Lowe VJ, Goodman PC, Herndon J. Thoracic nodal staging with positron emission tomography (PET) and ^{18}FDG in patients with bronchogenic carcinoma. *Chest* 1995; 108: 1617–1621.
3. Scott WJ, Gobar LS, Terry JD, Dewen NA, Sunderland JJ, Little AG. Mediastinal lymph node staging of non-small-cell lung cancer: A prospective comparison of computed tomography and positron emission tomography. *J Thorac Cardiovasc Surg* 1996; 111: 642–648.
4. Steinert HC, Hauser M, Allemann F, Engel H, Berthold T, von Schulthess GK, et al. Non-small cell lung cancer: Nodal staging with FDG PET versus CT with correlative lymph node mapping and sampling. *Radiology* 1997; 202: 441–446.
5. Guhlmann A, Strorck M, Kotzerke J, Moog F, Sunder-Plassmann L, Reske SN. Lymph node staging in non-small-cell lung cancer: evaluation by [^{18}F]FDG positron emission tomography (PET). *Thorax* 1997; 52: 438–441.
6. Boiselle PM, Patz EF Jr, Vining DJ, Weissleder R, Shephard JA, McLoud TC. Imaging of mediastinal lymph nodes: CT, MR, and FDG PET. *Radiographics* 1998; 18: 1061–1069.
7. Marom EM, McAdams HP, Erasmus JJ, Goodman PC, Culhane DK, Coleman RE. Staging non-small cell lung cancer with whole-body PET. *Radiology* 1999; 212: 803–809.
8. Ahuja V, Coleman RE, Herndon J, Patz EF Jr. The prognostic significance of fluorodeoxyglucose positron emission tomography imaging for patients with non-small cell lung carcinoma. *Cancer* 1998; 83: 918–924.
9. Shreve PD, Anzai Y, Wahl RL. Pitfalls in oncologic diagnosis with FDG PET imaging: Physiologic and benign variants. *Radiographics* 1999; 19: 61–77.
10. Bakheet SM, Sleem M, Powe J, Al Amro A, Larsson SG, Mahassin Z. F-18 fluorodeoxyglucose chest uptake in lung inflammation and infection. *Clin Nucl Med* 2000; 25: 273–278.
11. Inoue K, Matoba S. Counterattack of re-emerging tuberculosis after 38 years. *Int J Tuberc Lung Dis* 2001; 5: 873–875.
12. Fujimoto K, Edamitsu O, Meno S, Abe T, Honda N, Ogoh Y, et al. MR diagnosis for metastasis or non-metastasis of mediastinal and hilar lymph nodes in case of primary lung cancer: Detectability, signal intensity, and MR-pathologic correlation. *Nippon Igaku Hoshasen Gakkai Zasshi* 1995; 55: 162–171.
13. Vansteenkiste JF, Stroobants SG, De Leyn PR, Dupont PJ, Bogaert J, Maes A, et al. Lymph node staging in non-small-cell lung cancer with FDG-PET scan: A prospective study on 690 lymph node stations from 68 patients. *J Clin Oncol* 1998; 16: 2142–2419.
14. Liewald F, Grosse S, Storck M, Guhlmann A, Halter G, Reske S, et al. How useful is positron emission tomography for lymphnode staging in non-small-cell lung cancer? *Thorac Cardiovasc Surg* 2000; 48: 93–96.
15. Yoon YC, Lee KS, Shim YM, Kim B-T, Kim K, Kim TS. Metastasis to regional lymph nodes in patients with esophageal squamous cell carcinoma: CT versus FDG PET for presurgical detection—prospective study. *Radiology* 2003; 227: 764–770.
16. Fujii H, Yasuda S, Ide M, Takahashi W, Mochizuki Y, Shohtsu A, et al. Hilar activity on the F-18 FDG whole-body PET studies. *Jpn J Clin Radiol* 1999; 44: 199–206.
17. Shiraki N, Hara M, Ogino H, Shibamoto Y, Iida A, Tamaki T, et al. False-positive and true-negative hilar and mediastinal lymph nodes on FDG-PET—Radiological-pathological correlation—. *Ann Nucl Med* 2004; 18: 23–28.
18. Lowe VJ, Duhaylongsod FG, Patz EF Jr, Delong DM, Hoffman JM, Wolfe WG, et al. Pulmonary abnormalities and PET data analysis: A retrospective study. *Radiology* 1997; 202: 435–439.
19. Farrell MA, McAdams HP, Herndorn JE, Patz EF Jr. Non-small cell lung cancer: FDG PET for nodal staging in patients with stage I disease. *Radiology* 2000; 215: 886–890.
20. Kubota R, Yamada S, Kubota K, Ishiwata K, Tamahashi N, Ido T. Intratumoral distribution of fluorine-18-fluorodeoxyglucose *in vivo*: High accumulation in macrophages and granulation tissues studied by microautoradiography. *J Nucl Med* 1992; 33: 1972–1980.