N Stage Disease in Patients with Non–Small Cell Lung Cancer: Efficacy of Quantitative and Qualitative Assessment with STIR Turbo Spin-Echo Imaging, Diffusion-weighted MR Imaging, and Fluorodeoxyglucose PET/CT[1]

Yoshiharu Ohno, MD, PhD
Hisanobu Koyama, MD, PhD
Takeshi Yoshikawa, MD, PhD
Mizuho Nishio, MD
Nobukazu Aoyama, RT
Yumiko Onishi, MD, PhD
Daisuke Takenaka, MD
Sumiaki Matsumoto, MD, PhD
Yoshimasa Maniwa, MD, PhD
Wataru Nishio, MD
Yoshihiro Nishimura, MD, PhD
Tomoo Itoh, MD, PhD
Kazuro Sugimura, MD

Purpose: To prospectively compare the diagnostic capability of short inversion time inversion-recovery (STIR) turbo spin-echo (SE) imaging, diffusion-weighted (DW) magnetic resonance (MR) imaging, and fluorodeoxyglucose (FDG) combined positron emission tomography (PET) and computed tomography (CT) in N stage assessment in patients with non–small cell lung cancer (NSCLC).

Materials and Methods: This prospective study was approved by the institutional review board, and written informed consent was obtained from all patients. A total of 250 consecutive patients with NSCLC (136 men; mean age, 73 years; 114 women; mean age, 72 years) prospectively underwent pretherapeutic STIR turbo SE imaging, DW MR imaging, and FDG PET/CT, as well as surgical and pathologic examinations (N0 disease, $n = 157$; N1 disease, $n = 72$; N2 disease, $n = 16$; N3 disease, $n = 5$). Lymph node–to-saline ratio (LSR), lymph node–to-muscle ratio (LMR), apparent diffusion coefficient (ADC), maximal standardized uptake value (SUV_{max}), and visual scoring were assessed for 135 metastatic lymph nodes and 135 randomly selected nonmetastatic lymph nodes. Receiver operating characteristic curve analysis was used to determine feasible threshold values. Diagnostic capabilities for N stage assessment were compared with the McNemar test on a per-patient basis.

Results: When feasible, threshold values were used for quantitative assessment; sensitivity and accuracy of LSR and LMR (sensitivity, 82.8%; accuracy, 86.8%) proved to be significantly higher than those of ADC (sensitivity: 74.2%, $P = .01$; accuracy: 84.4%, $P = .04$) and SUV_{max} (sensitivity: 74.2%, $P = .01$). For qualitative assessment, sensitivity of STIR turbo SE imaging (77.4%) was significantly higher than that of DW MR imaging (71.0%, $P = .03$) and FDG PET/CT (69.9%, $P = .02$).

Conclusion: Quantitative and qualitative assessments of N stage disease in patients with NSCLC obtained with STIR turbo SE MR imaging are more sensitive and/or more accurate than those obtained with DW MR imaging and FDG PET/CT.

Supplemental material: *http://radiology.rsna.org/lookup/suppl/doi:10.1148/radiol.11110281/-/DC1*

[1] From the Department of Radiology (Y. Ohno, H.K., T.Y., M.N., Y. Onishi, D.T., S.M., K.S.) and Divisions of Thoracic Surgery (Y.M., W.N.), Respiratory Medicine (Y.N.), and Diagnostic Pathology (T.I.), Kobe University Graduate School of Medicine, 7-5-2 Kusunoki-cho, Chuo-ku, Kobe 650-0017, Japan; and Division of Radiology, Kobe University Hospital, Kobe, Japan (Y. Ohno, N.A.). Received February 8, 2011; revision requested March 22; revision received May 16; accepted June 21; final version accepted June 30. Supported by Philips Healthcare and a Grant-in-Aid for Scientific Research from the Japanese Ministry of Education, Culture, Sports, Science and Technology (JSTS.KAKEN no 20591442). **Address correspondence to** Y. Ohno (e-mail: *yosirad@med.kobe-u.ac.jp*).

Lung cancer is the leading cause of cancer death in both men and women worldwide (1). Non–small cell lung cancer (NSCLC) accounts for 80% of all cases of lung cancer, and it can be cured with surgery if it is detected at an early stage (2). In patients with NSCLC, involvement of mediastinal lymph nodes is an important prognostic factor because accurate staging of disease is needed to limit surgery or multimodality treatment to only those patients who might benefit from these treatments (3–5).

Since the early 1990s, fluorine 18 fluorodeoxyglucose (FDG) positron emission tomography (PET) and FDG PET combined with computed tomography (CT) have been found to enable more accurate evaluation of N stage disease in patients with NSCLC than use of computed tomography (CT) alone (6–11). However, the diagnostic capability of FDG PET and FDG PET/CT has been limited by the low spatial resolution of PET images, the many nodes with false-negative findings due to the difficulty of correctly interpreting lesions with high FDG uptake, and the nodes with false-positive findings, such as sarcoid reactions and acute or chronic inflammation (6–10). In routine clinical practice, many patients who are suspected of having lymph node metastases proceed directly to thoracotomy for primary resection combined with random hilar lymph node sampling. At other institutions, such patients routinely undergo mediastinoscopy before thoracotomy (12–14).

It was suggested in the past decade that short inversion time inversion-recovery (STIR) turbo spin-echo (SE) magnetic resonance (MR) imaging may be useful in distinguishing metastatic lymph nodes from nonmetastatic lymph nodes in patients with NSCLC with higher sensitivity and accuracy than is possible with CT, conventional T1-weighted MR imaging, FDG PET, or FDG PET/CT (15–18). As for STIR turbo SE imaging, the addition of increased T1 and T2 relaxation times yields a significantly higher net tissue contrast for malignant lymph nodes than for benign lymph nodes (15–18). In addition, the fat tissue signal is nulled or suppressed and considered useful in the detection of small lymph nodes (15–17). Furthermore, diffusion-weighted (DW) MR imaging has been put forward in the past few years as a new technique with which to differentiate metastatic lymph nodes from nonmetastatic lymph nodes in patients with NSCLC (19–21). DW MR imaging yields information, such as integrity, about microscopic structures (22). When apparent diffusion coefficients (ADCs) are assessed with DW MR imaging, there is a significant difference in ADC between malignant and benign lesions (19–21). To our knowledge, however, no one has directly compared the diagnostic capability for N stage assessment of STIR turbo SE imaging, DW MR imaging, and FDG PET/CT in a large prospective cohort.

We hypothesized that STIR turbo SE imaging might be more sensitive and more accurate than DW MR imaging and FDG PET/CT in the assessment of mediastinal and hilar lymph node metastases in patients with NSCLC. The purpose of this study was to prospectively compare the diagnostic capability of STIR turbo SE imaging, DW MR imaging, and FDG PET/CT in N stage assessment in patients with NSCLC.

Advances in Knowledge

- Per-patient analysis showed that quantitatively assessed short inversion time inversion-recovery (STIR) turbo spin-echo (SE) imaging (sensitivity, 82.8%; accuracy, 86.8%) was significantly more sensitive and more accurate in *N* stage assessment than diffusion-weighted (DW) MR imaging (sensitivity: 74.2%, $P = .01$; accuracy: 84.4%, $P = .04$) and fluorodeoxyglucose (FDG) PET/CT (sensitivity, 74.2%; $P = .01$).
- Qualitative assessment with STIR turbo SE imaging (sensitivity, 77.4%) was significantly more sensitive than qualitative assessment with DW MR imaging (71.0%, $P = .03$) and FDG PET/CT (69.9%, $P = .02$) on a per-patient basis.
- Lymph node–to-muscle ratio (LMR) (sensitivity, 82.8%; specificity, 89.2%; accuracy, 86.8%) may be substituted for lymph node–to-saline ratio (LSR) (sensitivity, 82.8%; specificity, 89.2%; accuracy, 86.8%) for quantitative N stage assessment with STIR turbo SE imaging in clinical practice.

Implication for Patient Care

- Comparison of quantitative and qualitative analyses of STIR turbo SE imaging, DW MR imaging, and integrated FDG PET/CT show that STIR turbo SE imaging had higher sensitivity (LSR, 82.8%; LMR, 82.8%; visual score, 77.4%) and accuracy (LSR, 86.8%; LMR, 86.8%) in N stage assessment than did the other modalities.

Materials and Methods

Study Protocol, Support, and Funding

This prospective study was approved by the institutional review board of Kobe

Published online before print
10.1148/radiol.11110281 **Content code:** CH

Radiology 2011; 261:605–615

Abbreviations:
ADC = apparent diffusion coefficient
DW = diffusion weighted
FDG = fluorodeoxyglucose
LMR = lymph node–to-muscle ratio
LSR = lymph node–to-saline ratio
NSCLC = non–small cell lung cancer
SE = spin echo
SI = signal intensity
STIR = short inversion time inversion recovery
SUV_{max} = maximum standard uptake value

Author contributions:
Guarantors of integrity of entire study, Y. Ohno, K.S.; study concepts/study design or data acquisition or data analysis/interpretation, all authors; manuscript drafting or manuscript revision for important intellectual content, all authors; manuscript final version approval, all authors; literature research, Y. Ohno; clinical studies, all authors; statistical analysis, Y. Ohno; and manuscript editing, Y. Ohno, K.S.

Potential conflicts of interest are listed at the end of this article.

University Hospital, and written informed consent was obtained from all patients. We received financial and/or technical supported from Philips Healthcare and a grant-in-aid for scientific research from the Japanese Ministry of Education, Culture, Sports, Science and Technology. None of the authors is an employee of Philips Healthcare. The authors had full control over the data for the duration of this study.

Subjects

Between January 2009 and December 2010, 250 consecutive patients with T1 or T2 NSCLC as evaluated on chest radiographs or CT images at a nearby hospital underwent contrast material–enhanced multidetector CT, STIR turbo SE imaging, DW MR imaging, FDG PET/CT, and mediastinoscopy before thoracotomy or resection of the primary lesion or before thoracotomy for primary resection in conjunction with hilar and mediastinal sampling. All patients were followed up for more than 1 year. This study group of 250 patients (mean age, 73 years; age range, 61–83 years) included 136 men (mean age, 73 years; age range, 61–80 years) and 114 women (mean age, 72 years; age range, 63–83 years).

The final diagnosis of lung cancer and N stage disease was based on pathologic findings in resected specimens, which showed that 218 patients had adenocarcinoma, 23 had squamous cell carcinoma, six had large cell carcinoma, and three had adenosquamous cell carcinoma. Stage N0 disease was detected in 157 patients, N1 disease was detected in 72, N2 disease was detected in 16, and N3 disease was detected in five.

MR Imaging

MR imaging was performed with a 1.5-T superconducting magnet (Gyroscan Achieva; Phillips Healthcare, Best, the Netherlands) by using a four-channel sensitivity encoding body coil. Axial and coronal STIR turbo SE images were obtained with a 0.9% saline phantom that was placed alongside the chest wall of the patient and by using a centrically reordered multishot black-blood STIR turbo SE sequence with sensitivity encoding (repetition time msec/effective echo time msec/inversion time msec, 1500–2000/8/150; echo train length, eight; two signals acquired; section thickness, 5 mm; section gap, 1.5 mm; 40–48 sections; 256 × 192 matrix; 512 × 384 reconstruction matrix; field of view, 320 mm; reduction factor, four). All DW MR imaging was performed by using the sequentially reordered half-Fourier single-shot STIR SE echo-planar imaging sequence (5000/70/180; echo train length, 41; section thickness, 5 mm; section gap, 1.5 mm; five signals acquired; *b* values, 0 and 1000 sec/mm^2; matrix, 96 × 96; reconstruction matrix, 256 × 256). Motion-probing gradients were added to the three axes (axial, sagittal, and coronal) on DW MR images. Respiratory triggering was not used, 20–24 15–16-second breath holds were required to obtain STIR turbo SE images, and DW MR images were obtained during free breathing. In each patient, mean acquisition time was 13.8 minutes (range, 12.0–15.0 minutes) for DW MR images and 6.0 minutes (range, 5.0–7.0 minutes) for STIR images.

Integrated FDG PET/CT

All patients fasted for at least 6 hours before intravenous administration of FDG at a rate of 3.3 MBq per kilogram of body weight, and images were obtained from the skull to the midthigh 60 minutes after completion of injection. All FDG PET/CT examinations were performed with a commercially available PET/CT scanner (Discovery ST; GE Healthcare, Milwaukee, Wis). The axes of the multidetector CT and PET systems were mechanically aligned so that the patient could be moved from the multidetector CT gantry to the PET gantry by simply changing the position of the examination table. The resultant PET and CT images were coregistered on hardware. CT was performed from the skull to the pelvic floor according to a standardized protocol with the following settings: 140 kV; 40 mA with automatic exposure control; tube rotation time, 0.6 second per rotation; detector collimation, 16 × 1.25 mm; beam pitch, 1.675; section thickness, 3.75 mm; and reconstruction pitch, 3.27 mm (to match PET section thickness). Patients maintained normal shallow respiration during CT image acquisition, and no iodinated contrast medium was administered. Immediately after unenhanced CT, PET was performed in the identical transverse field of view. The acquisition time was 2 minutes per table position.

All integrated PET/CT examinations were performed within 60 minutes. The CT data were resized from a 512 × 512 matrix to a 128 × 128 matrix to match the PET data so that the images could be fused and CT-based transmission maps could be generated. PET data sets were reconstructed iteratively with an ordered-subset expectation maximization algorithm and segmented attenuation correction (two iterations, 21 subsets) and with CT data. Coregistered images were displayed with commercially available software (Fusion Viewer; Nihon Medi-Physics, Nishinomiya, Japan).

Lymph Node Sampling, Histopathologic Examinations, and Metastatic and Nonmetastatic Lymph Node Selection

Cervical mediastinoscopy (supplemented by left anterior mediastinotomy if the tumor was in the left upper lobe) (n = 23) was selectively performed if the CT findings suggested invasion of the superior mediastinum or enlarged nodes (criteria will be described later). Two pulmonary surgeons (Y.M., W.N.) with 20 and 23 years of experience, respectively, performed complete standard mediastinal nodal sampling during mediastinoscopy.

During thoracotomy (n = 227), the same surgeons systematically performed ipsilateral hilar and mediastinal sampling at the locations specified by the regional lymph node classifications of the American Joint Committee on Cancer and the Union International Contre le Cancer (23–25). On the right side, lymph node sites 1 through 4 (highest mediastinal, upper paratracheal, prevascular and retrotracheal, and lower paratracheal, respectively) were dissected; the hilum was also dissected from posterior to inferior for the removal of lymph node sites 7 (subcarinal), 8 (paraesophageal), and 9 (pulmonary ligament) (23–25). On the left side, dissection at the aortic arch commenced when the

surgeon located the vagus nerve and its recurrent laryngeal branch and excised lymph node sites 4, 5 (subaortic), and 6 (para-aortic), after which dissection proceeded around the hilum as on the right side. The sites of these surgically dissected lymph nodes were matched to the lymph nodes identified on CT and MR images according to the American Joint Committee on Cancer and the Union International Contre le Cancer regional lymph node classification for lung cancer staging (23). Next, all resected lymph nodes were fixed in 10% buffered formalin and cut into 1-mm-thick slices after macrosopic examination by pathologists with more than 18 years of experience with chest diseases. To diagnose lymph nodes as metastatic or nonmetastatic, the same pathologists microscopically examined each lymph node in serial slices at a 1-mm interval.

In this study, 135 metastatic lymph nodes and 987 nonmetastatic lymph nodes were diagnosed. Of the 987 nonmetastatic lymph nodes, 135 were randomly and computationally selected with commercially available software (StatMate III; Atoms, Tokyo, Japan), and all 135 metastatic lymph nodes were included for per-node analyses. The 135 nonmetastatic lymph nodes comprised three lymphatic edemas; six coagulation necroses; 11 inflammatory nodes due to sarcoidosis, chronic interstitial lung disease resulting from collagen vascular diseases, chronic obstructive pulmonary disease, and secondary obstructive pneumonia; 13 fibrotic nodes; 17 silicotic nodes; 22 anthracosilicostic nodes; 28 calcified nodes; and 35 normal nodes.

Image Analysis

To more objectively compare the diagnostic accuracy of STIR turbo SE imaging, DW MR imaging, and integrated FDG PET/CT, we evaluated pretherapeutic lymph node–to-saline ratio (LSR) and lymph node–to-muscle ratio (LMR) at STIR turbo SE imaging, ADC at DW MR imaging, and maximum standard uptake value (SUV_{max}) at integrated FDG PET/CT. On the other hand, although visual assessments of signal intensity (SI) and FDG uptake within each lymph node are less objective, in routine clinical practice, these approaches may function as potential substitutes for LSR, LMR, ADC, and SUV_{max} when one distinguishes metastatic from nonmetastatic lymph nodes. Thus, pretherapeutic visual assessment of SI and FDG uptake within each lymph node were also performed, and their diagnostic accuracies were compared. All quantitative and qualitative assessments were performed with a commercially available picture archiving and communication system (Shade-Quest; Yokogawa, Tokyo, Japan). In addition, each investigator who performed quantitative or qualitative assessment with all modalities had no knowledge of the results of pathologic examination or any other investigation and evaluated all examinations in random order.

Quantitative Analysis of STIR Turbo SE, DW MR, and FDG PET/CT Images

On STIR turbo SE images, regions of interest were drawn by a chest radiologist with 17 years of experience (Y. Ohno) on the axial and coronal planes over 270 selected lymph nodes and the rhomboid muscle andaline phantom for measurement of SI. The regions of interest drawn over the lymph nodes encompassed the entire cross-sectional area of the lymph node (range, 16–576 mm^2), the saline phantom, and a rhomboid muscle of predetermined size (100 mm^2). The SI for each region of interest was determined as the mean obtained within the region of interest.

For traditional quantitative evaluation of SI of the lymph nodes, lymph node SI was normalized by comparing it wtih SI of 0.9% saline phantoms to yield the LSR. Previous studies (15–17) have shown an increase in LSR implies an increased likelihood of a metastatic lymph node, while a decrease in LSR implies an increased likelihood of a nonmetastatic lymph node.

LSR was determined with the following formula:

$$LSR = SI_{LN} \div SI_{SP}, \quad (1)$$

where SI_{LN} is the average SI of the lymph node and SI_{SP} is the average SI of the saline phantom.

For a new and easier way to normalize SI within lymph nodes observed on STIR turbo SE images, we calculated LMR with the following formula:

$$LMR = SI_{LN} \div SI_{RM}, \quad (2)$$

where SI_{RM} is the average SI of the rhomboid muscle.

For quantitative assessment of DW MR images of lymph nodes, the ADC was determined by using the region of interest measurement (range, 16–576 mm^2) and averaged to determine the final value of ADC for each lymph node. The same chest radiologist performed all region of interest placements on DW images. The ADC for each lymph node was calculated by using the following formula:

$$ADC = -[\ln(SI_h/SI_l)]/(bh-bl), \quad (3)$$

where SI_h and SI_l represent SI in the region of interest obtained with two gradient factors (bh and bl, respectively). For this study, bh was 1000 sec/mm^2, and bl was 0 sec/mm^2.

For quantitative assessment of lymph node metastases on integrated FDG PET/CT images, all SUV_{max} measurements were obtained from regions of interest drawn over mediastinal and hilar lymph nodes (range, 16–576 mm^2) on the axial and coronal planes by the same chest radiologist and averaged to determine the final value for each lymph node.

Qualitative Analysis of STIR Turbo SE, DW MR, and FDG PET/CT Images

To determine the qualitative analysis capability of STIR turbo SE imaging and DW MR imaging, all studies were prospectively interpreted by two chest radiologists with 5 and 9 years of experience (M.N., H.K.). On STIR turbo SE images, the probability of lymph node metastasis was assessed on a per-node basis by using the following five-point visual scoring system: 1, unable to evaluate; 2, equal to or less than mediastinal fat; 3, more than mediastinal fat and equal to or less than muscle; 4, more than muscle and equal to or less than primary lesion; and 5, more than

primary lesion. Each reviewer assessed all lymph nodes twice. The final diagnosis of lymph node metastasis was based on consensus of the two readers.

For qualitative assessment of DW MR images, the probability of lymph node metastasis was also assessed on a per-node basis by using the following five-point visual scoring system for DW MR images obtained with a *b* value of 1000 sec/mm^2: 1, unable to evaluate; 2, equal to or less than fat or skin; 3, more than fat or skin and equal to or less than muscle; 4, more than muscle and equal to or less than primary lesion; and 5, more than primary lesion. Each reviewer assessed all lymph nodes twice. The final probabiliy of lymph node metastasis was based on consensus of the two readers.

To assess the validity of qualitative analysis of FDG PET/CT images, two general radiologists with 8 and 14 years of experience (Y. Onishi and T.Y., respectively) prospectively evaluated all FDG PET/CT images. These two general radiologists also had more than 3 years of experience with PET. The probability of lymph node metastasis was evaluated on a per-node basis with the following five-point visual scoring system: 1, definitely absent; 2, probably absent; 3, equivocal; 4, probably present; and 5, definitely present. Each reviewer assessed all lymph nodes twice. The final probability of lymph node metastasis was again based on consensus of the two readers.

Statistical Analysis

To determine differences between metastatic and nonmetastatic lymph nodes in terms of quantitative indexes such as LSR, LMR, ADC, and SUV_{max}, these indexes were compared with the Student *t* test.

To assess observer performance for qualitatively analyzed STIR turbo SE imaging, DW MR imaging, and integrated FDG PET/CT on a per-node basis, a weighted κ statistic was used. Inter- and intraobserver agreement were considered slight if κ was less than 0.21, fair if κ was 0.21–0.40, moderate if κ was 0.41–0.60, substantial if κ was 0.61–0.80, and almost perfect if κ was 0.81–1.00 (26).

Receiver operating characteristic curve analyses were performed to compare the diagnostic capability of quantitative and qualitative assessments with STIR turbo SE imaging, DW MR imaging, and coregistered FDG PET/CT on a per-node basis. The feasible threshold value for each index was then determined and tested with the McNemar test; this test was selected because it enabled us to distinguish metastatic lymph nodes from nonmetastatic lymph nodes on a per-node basis. The feasible threshold value of each method was determined as demonstration of the highest accuracy, sensitivity, or both.

Weighted κ statistics were used to determine N stage agreement between pathologic examination findings and quantitative and qualitative assessments with each radiologic method. Agreement was considered slight if κ was less than 0.21, fair if κ was 0.21–0.40, moderate if κ was 0.41–0.60, substantial if κ was 0.61–0.80, and almost perfect if κ was 0.81–1.00 (26).

Finally, the McNemar test was used to compare sensitivity, specificity, and accuracy of *N* stage classifications as quantitatively and qualitatively assessed with STIR turbo SE imaging, DW MR imaging, and integrated FDG PET/CT.

For all statistical analyses, $P < .05$ was considered indicative of a significant difference.

Results

All STIR turbo SE imaging, DW MR imaging, and integrated FDG PET/CT procedures, as well as all pathologic examinations, were completed successfully. In this study, patients in whom N0 ($n = 157$) or N1 ($n = 72$) disease was diagnosed at pathologic analysis were assigned to the pathologic stage I or II group, respectively, and they underwent only thoracotomy. Patients in whom N2 ($n = 16$) or N3 ($n = 5$) disease was diagnosed at pathologic analysis were assigned to the pathologic stage IIIA or IIIB group, respectively. If pathologic stage IIIA ($n = 7$) or IIIB ($n = 5$) disease was diagnosed preoperatively, combined chemotherapy and radiation therapy was performed. On the other hand, when disease was classified as stage IIIA ($n = 9$) after thoracotomy, adjuvant chemotherapy was performed. Representative cases are shown in Figures 1 and 2.

Comparison of mean LSR, LMR, ADC, and SUV_{max} in metastatic (LSR: mean, 0.7 ± 0.1 [standard deviation]; range, 0.2–1.0; LMR: mean, 1.5 ± 0.3; range, 0.6–2.2; ADC: mean, [1.1 ± 0.9] × 10^{-3}sec/mm^2; range, [0.2–3.5] × 10^{-3}sec/mm^2; and SUV_{max}: mean, 2.2 ± 0.8; range, 0.2–4.5) and nonmetastatic (LSR: mean, 0.4 ± 0.2; range, 0.1–0.8; $P < .0001$; LMR: mean, 1.0 ± 0.3; range, 0.4–1.8; $P < .0001$; ADC: mean, [2.5 ± 0.8] × 10^{-3} sec/mm^2; range, [0.1–4.5] × 10^{-3} sec/mm^2; $P < .0001$; SUV_{max}: mean, 2.2 ± 0.8; range, 0.1–3.5; $P < .0001$) lymph nodes revealed significant differences.

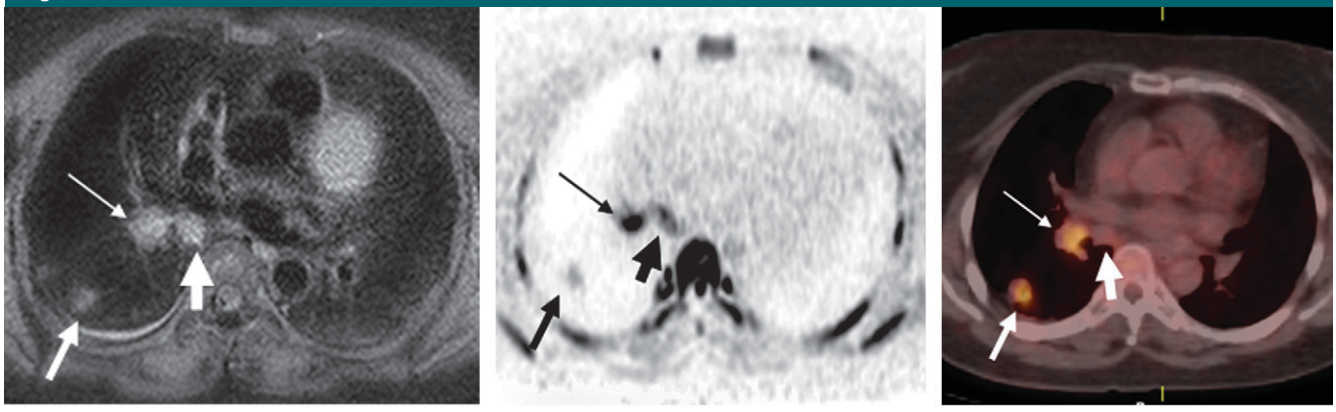
Details of intra- and interobserver agreements are shown in Tables E1 and E2 (online). Intraobserver agreement for STIR turbo SE imaging (κ = 0.91), DW MR imaging (κ = 0.87), and FDG PET/CT (κ = 0.90) was very good for one observer. In addition, intraobserver agreement for STIR turbo SE imaging (κ = 0.90), DW MR imaging (κ = 0.86), and FDG PET/CT (κ = 0.88) was very good in the other observer. On the other hand, interobserver agreement was very good for STIR turbo SE imaging (κ = 0.84) and FDG PET/CT (κ = 0.85) but only substantial for DW MR imaging (κ = 0.80).

Results of receiver operating characteristic curve analyses of all quantitative assessments performed with each method are shown in Figure 3. Areas under the receiver operating characteristic curve for STIR turbo SE imaging (LSR, 0.92; LMR, 0.89,), DW MR imaging (0.86), and FDG PET/CT (0.87) showed no significant differences.

Results of receiver operating characteristic curve analyses of all qualitative assessments performed with each method are shown in Figure 4. Areas under the receiver operating characteristic curve for STIR turbo SE imaging (0.87), DW MR imaging (0.84), and FDG PET/CT (0.86) also showed no significant differences.

Results of quantitative analyses of feasible threshold values, sensitivities,

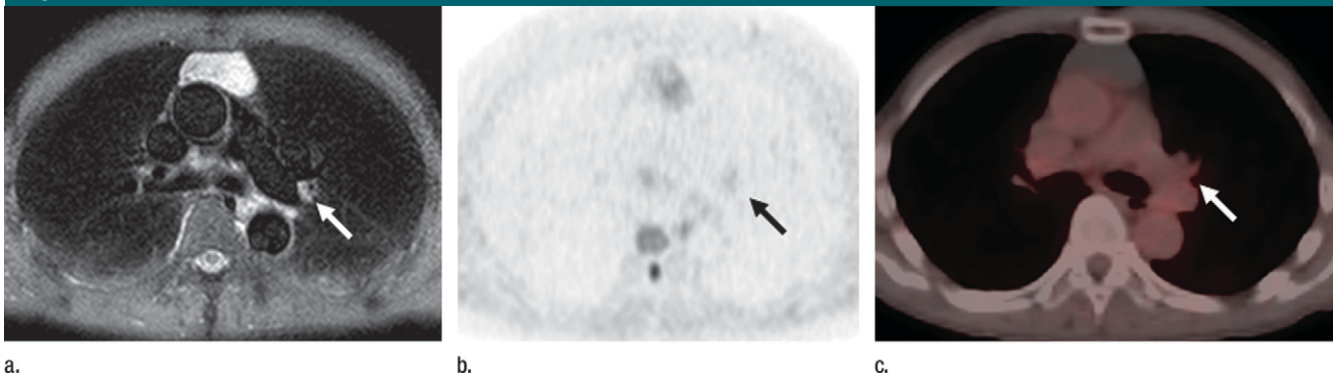
Figure 1



a. b. c.

Figure 1: Images in a 73-year old patient with pathologically diagnosed N2 adenocarcinoma. **(a)** STIR turbo SE image shows primary lesion (medium arrow), subcarina lymph node (thick arrow), and right hilar lymph node (thin arrow) have high SI. Primary lesion in the right lower lobe is visible in the same axial plane. LSRs of lymph nodes were 0.75 (right hilar lymph node) and 0.78 (subcarina lymph node), LMRs were 1.7 (right hilar lymph node) and 1.9 (subcarina lymph node), and visual scores were 5. An accurate diagnosis of N2 disease was made. **(b)** DW MR image shows primary lesion (medium arrow), subcarina lymph node (thick arrow), and right hilar lymph node (thin arrow) have high SI. Primary lesion in the right lower lobe is visible in the same axial plane. ADCs of lymph nodes were 2.8×10^{-3} sec/mm^2 (right hilar lymph node) and 3.4×10^{-3} sec/mm^2 (subcarina lymph node), and visual scores were 5. An accurate diagnosis of N2 disease was made. **(c)** FDG PET/CT image shows primary lesion (medium arrow) and right hilar lymph node (thin arrow) have high uptake of FDG and subcarina lymph node (thick arrow) has low uptake of FDG. Primary lesion in the right lower lobe is visible in the same axial plane. SUV_{max} of lymph nodes was 3.2 (right hilar lymph node) and 1.5 (subcarina lymph node), and visual scores were 5 (right hilar lymph node) and 2 (subcarina lymph node). An inaccurate diagnosis of N1 was made.

Figure 2



a. b. c.

Figure 2: Images in a 72-year old patient with pathologically diagnosed N1 adenocarcinoma. **(a)** STIR turbo SE image shows left hilar lymph node (arrow) has high SI. Primary lesion is not visible in the same axial plane. Thymic cyst can be seen in the anterior mediastinum. LSR of the lymph node was 0.70, LMR was 1.5, and visual score was 5. An accurate diagnosis of N1 disease was made. **(b)** DW MR image shows left hilar lymph node (arrow) has low SI. Primary lesion is not visible in the same axial plane. Thymic cyst can be seen as low SI in anterior mediastinum. ADC of the lymph node was 1.5×10^{-3} sec/mm^2, and visual score was 2. An inaccurate diagnosis of N0 was made. **(c)** FDG PET/CT image shows left hilar lymph node (arrow) has low uptake of FDG. Primary lesion is not visible in the same axial plane. Thymic cyst can be seen in the anterior mediastinum. SUV_{max} of the lymph node was 1.2, and visual score was 1. An inaccurate diagnosis of N0 disease was made.

specificities, positive and negative predictive values, and accuracies of STIR turbo SE imaging, DW MR imaging, and integrated FDG PET/CT on a per-node basis are shown in Table 1. When the feasible threshold values, which were defined as the highest accuracy and sensitivity in the differentiation of metastatic lymph nodes from nonmetastatic lymph nodes, for LSR, LMR, ADC, and SUV were applied, sensitivity and accuracy of LSR (sensitivity, 83.7%; $P = .0004$; accuracy, 85.1%; $P = .001$) and LMR (sensitivity, 81.5%; $P = .004$; accuracy, 83.7%; $P = .02$) were found to be significantly higher than those of ADC (sensitivity, 74.8%; accuracy, 81.1%). Sensitivity of LSR ($P = .001$) and LMR ($P = .008$) was significantly higher than that of SUV_{max} (75.6%). In addition, accuracy of LSR was significantly higher than that of SUV_{max} (accuracy, 82.2%; $P = .008$).

Results of qualitative analyses of feasible threshold values, which were

Figure 3

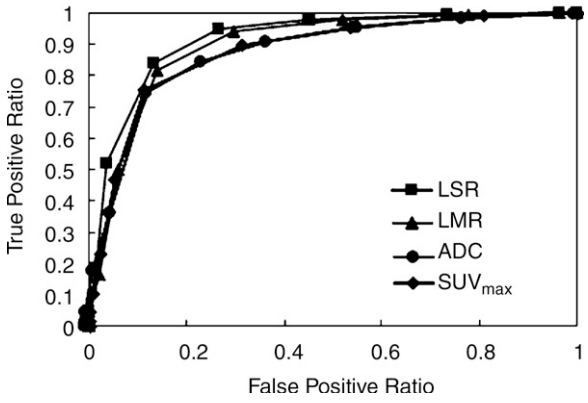


Figure 3: Receiver operating characteristic curve used for analysis of quantitative assessments with STIR turbo SE imaging, DW MR imaging, and integrated FDG PET/CT on a per-node basis. There were no significant differences between any of the quantitative methods. The threshold values were as follows: LSR, 0.6; LMR, 1.4; ADC, 2.5×10^{-3} sec/mm^2; and SUV_{max}, 2.0.

Figure 4

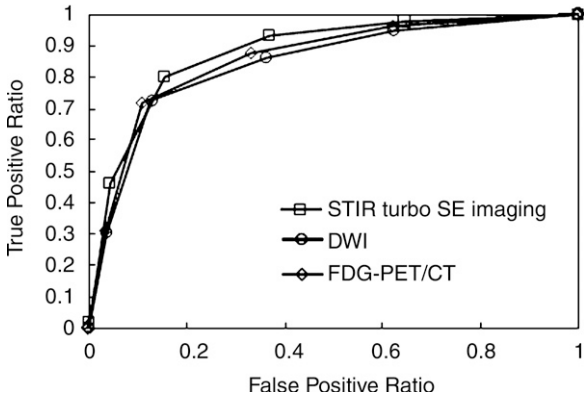


Figure 4: Receiver operating characteristic curve used for analysis of qualitative assessments with STIR turbo SE imaging, DW MR imaging, and integrated FDG PET/CT on a per-node basis. There were no significant differences between any of the qualitative methods. The threshold value for each method was 4.0.

determined as the highest accuracy for differentiation of metastatic from nonmetastatic lymph nodes, sensitivities, specificities, positive and negative predictive values, and accuracies of STIR turbo SE imaging, DW MR imaging, and integrated FDG PET/CT on a per-node basis are shown in Table 2. When the feasible threshold value for each visual scoring system was applied, sensitivity and accuracy of STIR turbo SE imaging (sensitivity, 80.0%; accuracy, 82.2%) proved to be significantly higher than those of DW MR imaging (sensitivity, 72.6%; P = .002; accuracy, 80.0%; P = .03). In addition, although sensitivity of STIR turbo SE imaging was significantly higher than that of FDG PET/CT (71.9%, P = .001), specificity of the former (84.4%) was significantly lower than that of the latter (88.9%, P = .03). In addition, although accuracy of LSR (85.1%) was significantly higher than that for qualitative assessment of STIR turbo SE imaging data (82.2%, P = .01), there was no significant difference between quantitative and qualitative assessment with each modality.

Details of agreements regarding N stage assessment among all methods and pathologic examination are shown in Tables E3–E5 (online). At assessment of agreements regarding quantitative N stage assessment among all methods and pathologic examination, κ values for LSR (κ = 0.77; 95% confidence interval: 0.69, 0.84), LMR (κ = 0.77; 95% confidence interval: 0.69, 0.84), ADC (κ = 0.72; 95% confidence interval: 0.64, 0.80), and SUV_{max} (κ = 0.75; 95% confidence interval: 0.67, 0.83) were all substantial. In addition, at assessment of agreements regarding qualitative N stage assessment among all methods and pathologic examination, κ values for STIR turbo SE imaging (κ = 0.74; 95% confidence interval: 0.66, 0.81), DW MR imaging (κ = 0.71; 95% confidence interval: 0.63, 0.80), and FDG PET/CT (κ = 0.71; 95% confidence interval: 0.62, 0.79) were also all substantial.

The results of quantitative comparative analyses of the diagnostic capabilities of N stage assessment among STIR turbo SE imaging, DW MR imaging, and integrated FDG PET/CT are shown in Table 3. When feasible threshold values were used for quantitative assessment, sensitivity and/or accuracy of LSR and LMR (sensitivity, 82.8%; accuracy, 86.8%) proved to be significantly higher than those of ADC (sensitivity, 74.2%; P = .01; accuracy, 84.4%; P = .04) and SUV_{max} (sensitivity, 74.2%; P = .01).

The results of qualitative comparative analyses of the diagnostic capabilities of N stage assessment with STIR turbo SE imaging, DW MR imaging, and integrated FDG PET/CT are shown in Table 4. When the feasible threshold values for qualitative assessment were applied, sensitivity of STIR turbo SE imaging (77.4%) was significantly higher than that of DW MR imaging (71.0%, P = .03) and FDG PET/CT (69.9%, P = .02). In addition, although accuracy of LSR (86.8%) and LMR (86.8%) was significantly higher than accuracy of qualitative assessment with STIR turbo SE imaging (84.4%, P = .04), there was no significant difference between quantitative and qualitative assessment for each modality.

Discussion

Our results showed that STIR turbo SE imaging is substantially more sensitive and accurate than DW MR imaging and integrated FDG PET/CT in quantitative and qualitative analyses of patients with NSCLC. In previous studies, N stage assessment with DW MR imaging was compared with N stage assessment with STIR turbo SE imaging and FDG PET/CT in a small cohort (18–21). However, to our knowledge, there have been no reports of direct comparison of diagnostic accuracy of STIR turbo SE imaging and DW MR imaging with use of 1.5-T MR systems with diagnostic accuracy of integrated FDG PET/CT in N stage

Table 1

Quantitative Assessments with STIR Turbo SE Imaging, DW MR Imaging, and FDG PET/CT on a Per-Node Basis

Modality and Measurement	Feasible Threshold Value	Sensitivity (%)	Specificity (%)	Positive Predictive Value (%)	Negative Predictive Value (%)	Accuracy (%)
STIR turbo SE imaging						
LSR	0.6	83.7 (113/135)	86.7 (117/135)	86.3 (113/131)	84.2 (117/139)	85.1 (230/270)
LMR	1.4	81.5 (110/135)	85.9 (116/135)	85.3 (110/129)	82.3 (116/141)	83.7 (226/270)
DW MR imaging						
ADC ($\times 10^{-3}$ sec/mm^2)	2.5	74.8 (101/135)*†	87.4 (118/135)	85.6 (101/118)	77.6 (118/152)	81.1 (219/270)*†
FDG PET/CT						
SUV_{max}	2.0	75.6 (102/135)*†	88.8 (120/135)	87.2 (102/117)	78.4 (120/153)	82.2 (222/270)*

Note.—Raw data are in parentheses.

* $P < .05$ when compared with LSR.

† $P < .05$ when compared with LMR.

Table 2

Qualitative Assessments with STIR Turbo SE Imaging, DW MR Imaging, and FDG PET/CT on a Per-Node Basis

Modality	Feasible Threshold Value	Sensitivity (%)	Specificity (%)	Positive Predictive Value (%)	Negative Predictive Value (%)	Accuracy (%)
STIR turbo SE imaging	4.0	80.0 (108/135)	84.4 (114/135)	83.7 (108/129)	80.9 (114/141)	82.2 (222/270)
DW MR imaging	4.0	72.6 (98/135)*	87.4 (118/135)	85.2 (98/115)	76.1 (118/155)	80.0 (216/270)*
FDG PET/CT	4.0	71.9 (97/135)*	88.9 (120/135)*	86.6 (97/112)	75.9 (120/158)	80.4 (217/270)

Note.—Raw data are in parentheses.

* $P < .05$ when compared with qualitatively assessed STIR turbo SE imaging.

assessment in patients with NSCLC in a large prospective cohort. According to our results, in 250 consecutive patients with NSCLC, STIR turbo SE imaging would perform better than DW MR imaging in the differentiation of metastatic lymph nodes from nonmetastatic lymph nodes and in N stage assessment.

Comparison of LSR, LMR, and SUV_{max} assessment of each quantitative index for metastatic and nonmetastatic lymph nodes showed that all indexes of metastatic lymph nodes were significantly higher than all indexes of nonmetastatic lymph nodes, and the findings of this comparison were considered compatible with those reported in the literature (6–19). On the other hand, ADC of metastatic lymph nodes was significantly lower than ADC of nonmetastatic lymph nodes because of higher cell density, and ADC of metastatic lymph nodes was considered compatible with ADC values reported in the literature (19–21). However, highly measured ADC values of nonmetastatic lymph nodes might be contaminated by necrosis, perfusion, and motion artifacts because we adapted a *b* value of 0 sec/mm^2 at DW MR imaging.

Intraobserver agreement for both observers for STIR turbo SE imaging, DW MR imaging, and FDG PET/CT was almost perfect. In addition, interobserver agreement for visual assessment of STIR turbo SE and FDG PET/CT images was almost perfect. Interobserver agreement for visual assessment of DW MR images was substantial. These results suggested that visual assessment of DW MR images was less reproducible than visual assessment of STIR turbo SE and FDG PET/CT images. Lower interobserver agreement for DW MR imaging might be affected by lower spatial resolution; less anatomic information due to no coregistration of DW MR images with superiorly visualized anatomic information, such as CT or T1-weighted MR images; lower image quality due to no adaptation of motion correction; and overlapped SI between metastatic and nonmetastatic lymph nodes. Thus, improvement in the diagnostic performance of DW MR imaging is important if this technique is to be used in N stage assessment in routine clinical practice.

Direct comparison of quantitatively and qualitatively assessed data obtained with STIR turbo SE imaging, DW MR imaging, and FDG PET/CT on a per-node basis revealed that STIR turbo SE imaging should be considered more sensitive and/or more accurate than DW MR imaging or FDG PET/CT, both quantitatively and qualitatively, on a per-node basis. In addition, no significant differences were observed between DW MR imaging and FDG PET/CT. This finding suggests that DW MR imaging has potential as a substitute for FDG PET/CT on a per-node basis. These results are similar to those reported in the literature (15–21). In addition, false-positive nodules on STIR turbo SE images and DW MR images were lymphatic edemas, coagulation necroses, and inflammatory nodes. On the other hand, false-positive nodules on FDG PET/CT images consisted of inflammatory and anthracosilicostic nodes.

Table 3

Comparison of Quantitative Diagnostic Capability in N Stage Assessment of STIR Turbo SE Imaging, DW MR Imaging, and FDG PET/CT

Modality	Sensitivity (%)	Specificity (%)	Positive Predictive Value (%)	Negative Predictive Value (%)	Accuracy (%)
STIR turbo SE imaging					
LSR	82.8 (77/93)	89.2 (140/157)	81.9 (77/94)	89.7 (140/156)	86.8 (217/250)
LMR	82.8 (77/93)	89.2 (140/157)	81.9 (77/94)	89.7 (140/156)	86.8 (217/250)
DW MR imaging					
ADC ($\times 10^{-3}$ sec/mm^2)	74.2 (69/93)*†	90.4 (142/157)	82.1 (69/84)	85.5 (142/166)	84.4 (211/250)*†
FDG PET/CT					
SUV_{max}	74.2 (69/93)*†	92.4 (145/157)	85.2 (69/81)	85.8 (145/169)	85.6 (214/250)

Note.—Raw data are in parentheses.

* $P < .05$ when compared with LSR.

† $P < .05$ when compared with LMR.

Table 4

Comparison of Qualitative Diagnostic Capability in N Stage Assessment of STIR Turbo SE Imaging, DW MR Imaging, and FDG PET/CT

Modality	Sensitivity (%)	Specificity (%)	Positive Predictive Value (%)	Negative Predictive Value (%)	Accuracy (%)
STIR turbo SE imaging	77.4 (72/93)	88.5 (139/157)	80.0 (72/90)	86.9 (139/160)	84.4 (211/250)
DW MR imaging	71.0 (66/93)*	89.8 (141/157)	80.5 (66/82)	83.9 (141/168)	82.8 (207/250)
FDG PET/CT	69.9 (65/93)*	91.7 (144/157)	83.3 (65/78)	83.7 (144/172)	83.6 (209/250)

Note.—Raw data are in parentheses.

* $P < .05$ when compared with qualitatively assessed STIR turbo SE imaging.

Our direct assessment of agreement regarding N stage determined with the different methods suggested that each method demonstrated substantial agreement with pathologic findings. However, quantitative and qualitative assessments with DW MR imaging and FDG PET/CT showed a tendency to underestimate N2 or N3 disease compared with assessments performed with STIR turbo SE imaging. These results suggest that reviewers must have missed more mediastinal lymph node metastases with DW MR imaging, FDG PET/CT, or both than they missed with STIR turbo SE imaging; this was because of the lower spatial resolution, lower contrast resolution between metastatic and nonmetastatic lymph nodes, and the presence of nonsynchronized cardiac and respiratory motion artifacts. In addition, DW MR images were obtained with the SE echo-planar sequence, which is more sensitive to susceptibility artifacts within the lung than is the turbo SE sequence. Thus, the previously mentioned sequential problems might be affected by underestimation of N stage at DW MR imaging as compared with N stage at pathologic examination.

The comparison of diagnostic capabilities of STIR turbo SE imaging, DW MR imaging, and integrated FDG PET/CT in N stage assessment revealed that STIR turbo SE imaging was more sensitive and/or more accurate than DW MR imaging or FDG PET/CT in quantitative and qualitative analyses. These results are compatible with those previously reported (15–21). In addition, the higher sensitivity and/or accuracy of STIR turbo SE imaging may well be the main reason for adoption of this technique over DW MR imaging or FDG PET/CT in surgical treatment or lymph node sampling during thoracotomy or mediastinoscopy before treatment. This is because accurate pathologic TNM staging after surgical treatment or before chemotherapy, radiation therapy, or both is essential to an accurate prognosis. In addition, LSR and LMR were not significantly different in terms of diagnostic performance. Thus, LMR, which shows SI of the lymph node normalized by that of muscle within the same imaging plane, is considered adequate for quantitative assessment of STIR turbo SE images in routine clinical practice.

There were limitations to our study. Although we used hematoxylin-eosin stains to histologically examine all lymph nodes at every 1 or 2 mm of section thickness, some microsopic metastatic nodes might have been missed (27,28). Thus, the pathologic diagnosis of lymph node metastases might have been affected. A cost analysis would be warranted to enable a more realistic determination of the practical significance of STIR turbo SE imaging in the diagnosis of metastatic lymph nodes. In addition, the probability of lymph node metastases was quantitatively and qualitatively evaluated at STIR turbo SE imaging, DW MR imaging, and FDG PET/CT in this study. However, we compared diagnostic capability of each evaluation among these three modalities,

and we did not assess diagnostic capability of qualitative assessment of each modality combined with quantitative measurement. When one considers the routine clinical practice, it would be helpful to evaluate diagnostic capability of qualitative assessment of each modality combined with quantitative measurement. Thus, it would be advisable to conduct both a cost-effectiveness analysis to reveal the diagnostic capability of qualitative assessment of each modality combined with quantitative measurement and a multicenter study with a larger prospective cohort in the near future to determine the true value of STIR turbo SE imaging in the care of patients with NSCLC.

In conclusion, quantitative and qualitative N stage assessment of patients with NSCLC performed with STIR turbo SE imaging is more sensitive and/or more accurate than assessment performed with DW MR imaging or FDG PET/CT. Thus, STIR turbo SE imaging may be used before surgical treatment or lymph node sampling during thoracotomy or mediastinoscopy in accurate pathologic TNM staging after surgical treatment or before chemotherapy, radiation therapy, or both. In addition, LMR may be substituted for LSR for use with quantitative N stage assessment with STIR turbo SE imaging in clinical practice.

Acknowledgments: The authors thank Kazuyuki Kobayashi, MD, PhD; Yasuhiro Funada, MD, PhD; and Yoshikazu Kotani, MD, PhD (Division of Cardiovascular and Respiratory Medicine, Department of Internal Medicine, Kobe University Graduate School of Medicine) and Chiho Ohbayashi, MD, and Yasuhiro Sakai, MD (Department of Diagnostic Pathology, Kobe University Graduate School of Medicine), for their contributions to this work. The authors also thank Yoshiyuki Ohno, MD, PhD, MPH, Professor Emeritus, Nagoya University (Department of Preventive Medicine, Graduate School of Medicine), for his statistical advice for completion of this manuscript.

Disclosures of Potential Conflicts of Interest: **Y. Ohno** Financial activities related to the present article: received a research grant from Philips Healthcare. Financial activities not related to the present article: none to disclose. Other relationships: none to disclose. **H.K.** No potential conflicts of interest to disclose. **T.Y.** No potential conflicts of interest to disclose. **M.N.** No potential conflicts of interest to disclose. **N.A.** No potential conflicts of interest to disclose. **Y. Onishi** No potential conflicts of interest to disclose. **D.T.** No potential conflicts of interest to disclose. **S.M.** No potential conflicts of interest to disclose. **Y.M.** No potential conflicts of interest to disclose. **W.N.** No potential conflicts of interest to disclose. **Y.N.** No potential conflicts of interest to disclose. **T.I.** No potential conflicts of interest to disclose. **K.S.** Financial activities related to the present article: received a research grant from Philips Healthcare. Financial activities not related to the present article: none to disclose. Other relationships: none to disclose.

References

1. Jemal A, Chu KC, Tarone RE. Recent trends in lung cancer mortality in the United States. J Natl Cancer Inst 2001;93(4):277–283.
2. Free CM, Ellis M, Beggs L, Beggs D, Morgan SA, Baldwin DR. Lung cancer outcomes at a UK cancer unit between 1998-2001. Lung Cancer 2007;57(2):222–228.
3. Okada M, Yoshikawa K, Hatta T, Tsubota N. Is segmentectomy with lymph node assessment an alternative to lobectomy for nonsmall cell lung cancer of 2 cm or smaller? Ann Thorac Surg 2001;71(3):956–960; discussion 961.
4. Onishi H, Araki T, Shirato H, et al. Stereotactic hypofractionated high-dose irradiation for stage I nonsmall cell lung carcinoma: clinical outcomes in 245 subjects in a Japanese multiinstitutional study. Cancer 2004;101(7):1623–1631.
5. Nakamura K, Saji H, Nakajima R, et al. A phase III randomized trial of lobectomy versus limited resection for small-sized peripheral non-small cell lung cancer (JCOG0802/WJOG4607L). Jpn J Clin Oncol 2010;40(3):271–274.
6. Wahl RL, Quint LE, Greenough RL, Meyer CR, White RI, Orringer MB. Staging of mediastinal non-small cell lung cancer with FDG PET, CT, and fusion images: preliminary prospective evaluation. Radiology 1994; 191(2):371–377.
7. Marom EM, McAdams HP, Erasmus JJ, et al. Staging non-small cell lung cancer with whole-body PET. Radiology 1999;212(3):803–809.
8. Antoch G, Stattaus J, Nemat AT, et al. Non-small cell lung cancer: dual-modality PET/CT in preoperative staging. Radiology 2003;229(2):526–533.
9. Shim SS, Lee KS, Kim BT, et al. Non-small cell lung cancer: prospective comparison of integrated FDG PET/CT and CT alone for preoperative staging. Radiology 2005;236(3):1011–1019.
10. Halpern BS, Schiepers C, Weber WA, et al. Presurgical staging of non-small cell lung cancer: positron emission tomography, integrated positron emission tomography/CT, and software image fusion. Chest 2005;128(4): 2289–2297.
11. Silvestri GA, Gould MK, Margolis ML, et al. Noninvasive staging of non-small cell lung cancer: ACCP evidenced-based clinical practice guidelines (2nd edition). Chest 2007; 132(3 Suppl):178S–201S.
12. Gdeedo A, Van Schil P, Corthouts B, Van Mieghem F, Van Meerbeeck J, Van Marck E. Prospective evaluation of computed tomography and mediastinoscopy in mediastinal lymph node staging. Eur Respir J 1997;10(7): 1547–1551.
13. Deslauriers J, Grégoire J. Clinical and surgical staging of non-small cell lung cancer. Chest 2000;117(4 Suppl 1):96S–103S.
14. Detterbeck FC, Jantz MA, Wallace M, Vansteenkiste J, Silvestri GA; American College of Chest Physicians. Invasive mediastinal staging of lung cancer: ACCP evidence-based clinical practice guidelines (2nd edition). Chest 2007;132(3 Suppl): 202S–220S.
15. Takenaka D, Ohno Y, Hatabu H, et al. Differentiation of metastatic versus nonmetastatic mediastinal lymph nodes in patients with non-small cell lung cancer using respiratory-triggered short inversion time inversion recovery (STIR) turbo spin-echo MR imaging. Eur J Radiol 2002;44(3):216–224.
16. Ohno Y, Hatabu H, Takenaka D, et al. Metastases in mediastinal and hilar lymph nodes in patients with non-small cell lung cancer: quantitative and qualitative assessment with STIR turbo spin-echo MR imaging. Radiology 2004;231(3):872–879.
17. Ohno Y, Koyama H, Nogami M, et al. STIR turbo SE MR imaging vs coregistered FDG-PET/CT: quantitative and qualitative assessment of N-stage in non-small-cell lung cancer patients. J Magn Reson Imaging 2007;26(4):1071–1080.
18. Morikawa M, Demura Y, Ishizaki T, et al. The effectiveness of 18F-FDG PET/CT combined with STIR MRI for diagnosing nodal involvement in the thorax. J Nucl Med 2009; 50(1):81–87.
19. Nomori H, Mori T, Ikeda K, et al. Diffusion-weighted magnetic resonance imaging can be used in place of positron emission tomography for N staging of non-small cell lung cancer with fewer false-positive results. J Thorac Cardiovasc Surg 2008;135(4):816–822.
20. Hasegawa I, Boiselle PM, Kuwabara K, Sawafuji M, Sugiura H. Mediastinal lymph nodes in patients with non-small cell lung cancer: preliminary experience with

diffusion-weighted MR imaging. J Thorac Imaging 2008;23(3):157–161.

21. Nakayama J, Miyasaka K, Omatsu T, et al. Metastases in mediastinal and hilar lymph nodes in patients with non-small cell lung cancer: quantitative assessment with diffusion-weighted magnetic resonance imaging and apparent diffusion coefficient. J Comput Assist Tomogr 2010;34(1):1–8.
22. Koh DM, Collins DJ. Diffusion-weighted MRI in the body: applications and challenges in oncology. AJR Am J Roentgenol 2007;188(6):1622–1635.
23. Cymbalista M, Waysberg A, Zacharias C, et al. CT demonstration of the 1996 AJCC-UICC regional lymph node classification for lung cancer staging. RadioGraphics 1999;19(4):899–900.
24. Graham AN, Chan KJ, Pastorino U, Goldstraw P. Systematic nodal dissection in the intrathoracic staging of patients with non-small cell lung cancer. J Thorac Cardiovasc Surg 1999;117(2):246–251.
25. Watanabe S, Oda M, Tsunezuka Y, Go T, Ohta Y, Watanabe G. Peripheral small-sized (2 cm or less) non-small cell lung cancer with mediastinal lymph node metastasis; clinicopathologic features and patterns of nodal spread. Eur J Cardiothorac Surg 2002;22(6):995–999.
26. Ludbrook J. Statistical techniques for comparing measurers and methods of measurement: a critical review. Clin Exp Pharmacol Physiol 2002;29(7):527–536.
27. Cai J, Ikeguchi M, Tsujitani S, Maeta M, Liu J, Kaibara N. Significant correlation between micrometastasis in the lymph nodes and reduced expression of E-cadherin in early gastric cancer. Gastric Cancer 2001;4(2):66–74.
28. Okada Y, Fujiwara Y, Yamamoto H, et al. Genetic detection of lymph node micrometastases in patients with gastric carcinoma by multiple-marker reverse transcriptase-polymerase chain reaction assay. Cancer 2001;92(8):2056–2064.