

Preoperative staging of non-small-cell lung cancer: comparison of whole-body diffusion-weighted magnetic resonance imaging and ^{18}F -fluorodeoxyglucose-positron emission tomography/computed tomography

Gregor Sommer · Mark Wiese · Leopold Winter ·
Claudia Lenz · Markus Klarhöfer · Flavio Forrer ·
Didier Lardinois · Jens Bremerich

Received: 13 January 2012 / Revised: 18 April 2012 / Accepted: 11 May 2012 / Published online: 9 July 2012
© European Society of Radiology 2012

Abstract

Objective To investigate the diagnostic value of whole-body magnetic resonance imaging (MRI) including diffusion-weighted imaging with background signal suppression (DWIBS) for preoperative assessment of non-small-cell lung cancer (NSCLC) in comparison to ^{18}F -fluorodeoxyglucose (^{18}F FDG) positron emission tomography/computed tomography (PET/CT).

Methods Thirty-three patients with suspected NSCLC were enrolled. Patients were examined before surgery with PET/CT and whole-body MRI including T1-weighted turbo spin echo (TSE), T2-weighted short tau inversion recovery (STIR) and DWIBS sequences ($b=0/800$). Histological or cytological specimens were taken as standard of reference.

Results Whole-body MRI with DWIBS as well as PET/CT provided diagnostic image quality in all cases. Sensitivity for primary tumour detection: MRI 93%, PET/CT 98%. T-

staging accuracy: MRI 63%, PET/CT 56%. N-staging accuracy: MRI 66%, PET/CT 71%. UICC staging accuracy: MRI 66%, PET/CT 74%. Sensitivity for metastatic involvement of individual lymph node groups: MRI 44%, PET/CT 47%. Specificity for individual non-metastatic lymph node groups: MRI 93%, PET/CT 96%. Assessment accuracy for individual lymph node groups: MRI 85%, PET/CT 88%. Observer agreement rate for UICC staging: MRI 74%, PET/CT 90%.

Conclusion Whole-body MRI with DWIBS provides comparable results to PET/CT in staging of NSCLC, but shows no superiority. Most relevant challenges for both techniques are T-staging accuracy and sensitivity for metastatic lymph node involvement.

Key Points

- Numerous radiological methods are available for the crucial staging of lung cancer
- Whole-body DWIBS MRI provides comparable results to PET/CT in NSCLC staging.
- No evident superiority of whole-body DWIBS over PET/CT in NSCLC staging.
- Challenges for both techniques are T-staging and detection of small metastases.

G. Sommer (✉) · L. Winter · F. Forrer · J. Bremerich
Clinic of Radiology and Nuclear Medicine,
University of Basel Hospital,
Petersgraben 4,
4031 Basel, Switzerland
e-mail: gsommer@uhbs.ch

M. Wiese · D. Lardinois
Clinic of Thoracic Surgery, University of Basel Hospital,
Spitalstrasse 21,
4031 Basel, Switzerland

C. Lenz · M. Klarhöfer
Division of Radiological Physics, Clinic of Radiology
and Nuclear Medicine, University of Basel Hospital,
Petersgraben 4,
4031 Basel, Switzerland

M. Klarhöfer
Siemens Switzerland Ltd., Healthcare Sector,
Zurich, Switzerland

Keywords Non-small-cell lung carcinoma · Neoplasm staging · Whole-body imaging · Diffusion magnetic resonance imaging · Positron emission tomography

Introduction

Lung cancer is the leading cause of cancer-related death and accounts for 14% of all new cases of cancer in the USA in 2011 [1]. Accurate staging is mandatory to select the most appropriate therapy and to determine prognosis. Combined

^{18}F -fluorodeoxyglucose positron emission tomography and computed tomography (PET/CT) is considered as the standard of reference for preoperative assessment of non-small-cell lung cancer (NSCLC) [2, 3]. Despite its widespread use and high degree of standardisation, the results provided by PET/CT are still not totally satisfying. Limitations of PET/CT are particularly limited spatial resolution [4] and low specificity in distinguishing malignant lymphadenopathy from inflammatory changes, resulting in a considerable number of false-positive findings [5]. Moreover, PET/CT is associated with a considerable radiation burden to patients and medical personnel.

Magnetic resonance imaging (MRI) is currently the only technique that enables non-invasive whole-body assessment without ionising radiation. Another strength of MRI is its capability to create high soft tissue contrast without external contrast agents and with high spatial resolution. A novel powerful source of contrast generation for whole-body MRI became available in 2004 with the introduction of diffusion-weighted imaging with background signal suppression (DWIBS) [6]. The image contrast of DWIBS is based on the diffusion properties of water molecules and reflects tissue parameters like cellular density and tissue architecture [7]. In the last few years, DWIBS has been investigated successfully in many fields of oncology [8]. Also for NSCLC, an increasing number of studies on DWI became available [9–24]. Only a few studies have addressed validation DWI versus PET/CT as standard of reference. Among the data that have been published up till now, four studies focus on tumour detection and characterisation of lung nodules [9–12], four studies on N-staging and characterisation of mediastinal lymph nodes [11–14], and two studies on M-staging [12, 15]. Only one study [11] provides data on overall staging accuracy according to the Union for International Cancer Control (UICC) classification. To the best of our knowledge, this is the first study presenting data on T-staging accuracy of MRI with diffusion weighting.

The objective of this study is to assess the diagnostic value of whole-body MRI with DWIBS in comparison to PET/CT for comprehensive preoperative assessment of NSCLC in a clinical setting. Data evaluation included primary tumour detection, T-staging, detection of individual lymph node metastases, N-staging and UICC staging with histopathology and cytology as the reference standards.

Materials and methods

Patients

Thirty-three patients (24 men, 9 women, mean age 63.7 years, median age 66 years, age range 43–83 years) with suspected NSCLC were prospectively enrolled. All patients underwent

PET/CT and were scheduled for surgery according to the PET/CT findings. Whole-body MRI examinations were performed before surgery. In one patient, surgery was cancelled due to negative results in an additional transbronchial biopsy that was done after the MRI exam. The histological subtypes of pulmonary malignancies represented according to the WHO/IASLC classification [25] were adenocarcinoma (16 cases), squamous cell carcinoma (eight cases), adenosquamous carcinoma (one case), large cell carcinoma (two cases) and well differentiated neuroendocrine tumour (one case). Two patients were excluded from statistical evaluation as histology revealed malignant lesions of non-pulmonary origin (one lymphoma, one colon cancer metastasis). Three patients were diagnosed as negative for lung cancer by histology or cytology (one tuberculosis, one respiratory bronchiolitis-interstitial lung disease [RB-ILD], one inconclusive). The mean time interval between imaging and surgery was 34 (\pm 26) days for PET/CT and 3 (\pm 2) days for MRI. Three patients received surgical resection after neo-adjuvant radiation and/or chemotherapy. However, no therapy was performed between imaging and surgery. All procedures were in accordance with the ethical standards of the World Medical Association and written informed consent was obtained from all patients. The study was approved by the local ethics committee.

Imaging protocol

PET/CT examinations were performed on an integrated PET/CT system with 16-slice CT (Discovery; GE Healthcare, Chalfont St Giles, UK). ^{18}F -FDG was administered in a standard dose of 5 MBq per kg body weight (maximum dose 500 MBq) 60 min before imaging after a fasting period of a minimum of 6 h. For PET acquisition, eight bed positions with each 4-min data acquisition were obtained from skull to upper thigh. All patients received unenhanced low-dose CT for attenuation correction and anatomical reference (tube voltage=120 kV, tube current=100 mA, collimation=16 \times 3.75 mm, free breathing). In patients who had not previously undergone a dedicated chest CT examination, additional contrast enhanced CT of the chest was performed (tube voltage=120 kV, tube current=300 mA, collimation=16 \times 1.25 mm, breath hold, 80 ml intravenous iodinated contrast medium).

MRI examinations were performed on a 1.5-T whole-body MRI (Magnetom Avanto or Magnetom Symphony, Siemens Healthcare, Erlangen, Germany) using a dedicated 18-channel coil array system (total imaging matrix [Tim], Siemens Healthcare). The sequences employed were T1-weighted turbo spin echo (TSE) (TR=682 ms, TE=11 ms, matrix size=320 \times 240 pixels, slice thickness=5 mm, field of view (FoV)=500 \times 375 mm², acquisition time 7 \times 1:01 min), T2-weighted short tau inversion recovery (STIR) (TR=9,930 ms, TE=86 ms, TI=160 ms, matrix size=320 \times 240 pixels, slice thickness=5 mm, FoV=500 \times 375 mm²,

acquisition time $7 \times 1:19$ min), and DWIBS (single shot echo planar imaging [ss-EPI], TR=5,400 ms, TE=58 ms, b=0 and 800 s/mm², STIR fat suppression with TI=180 ms, matrix size=192×144 pixels, slice thickness=5 mm, FoV=500×375 mm², four averages, acquisition time $7 \times 1:43$ min) in transverse orientation, each covering the patients' body in seven acquisition steps from skull to upper thigh. Total examination time was 30 min. All data were acquired during free breathing. No contrast agent was applied. Apparent diffusion coefficients (ADCs) were calculated pixel-wise (linear fit to logarithmical data) and displayed as ADC maps. Image fusion of T1-weighted and high b-value DWI data was performed three-dimensionally in a semiautomatic fashion and rendered in transverse and coronal image stacks.

Image analysis

Initial assessment of the PET/CT examinations was done as part of the routine work by a board-certified nuclear medicine physician with more than 5 years' experience in PET/CT reading (Reader 1). Separate blinded study readings were performed >4 weeks later in three cases where data from a previous PET/CT examination had been used during the initial reading procedures. Second reading of all PET/CT examinations was done by a board-certified nuclear medicine physician with more than 10 years' experience in PET/CT reading (Reader 2). Both low-dose and contrast enhanced CT datasets were used for the PET/CT readings. MRI examinations were read by a board-certified radiologist with more than 5 years' experience in MRI reading (Reader 1) and in a second reading procedure by a board-certified radiologist with more than 10 years' experience in MRI reading (Reader 2). All datasets (T1-weighted, T2-weighted STIR and DWIBS) were considered for diagnosis. All readers were blinded to the results of the other imaging technique and histopathology results. Previous images from investigations other than PET/CT and MRI were available to all readers. Image reading was done on commercially available workstations (Centricity RA 1000; GE Healthcare).

Image quality of MRI was assessed on a four-point scale (very good = dataset with no visible artefacts; good = slight motion artefacts that do not affect diagnostic assessment; fair = artefacts that tolerably affect diagnostic assessment; unsatisfactory = non-diagnostic data). Evaluation was done both for the entire study and separately for the three individual sequences. Image quality of PET/CT was assessed overall using the same scale as for the MRI examinations.

Staging was done for each technique according to the 7th edition TNM and UICC classifications [26]. Both PET/CT and MRI images were interpreted in a qualitative manner considering both morphological and functional information. Increased FDG-uptake and restricted diffusion were identified by visual comparison of the lesion's signal to the FDG-

uptake of the liver parenchyma in PET and the background signal in high b-value DWI, respectively. Quantitative values for standardised uptake value (SUV) and ADC were calculated and used for interpretation of particular findings if found appropriate by the readers. However, no general cut-off values were applied for differentiating benign from malignant lesions. For lymph node assessment, a short axis diameter of >1 cm was regarded as a morphological criterion for metastatic involvement.

Lymph node stations were divided into three groups for individual assessment (S_n =stations according to the IASLC lymph node map [27], i = ipsilateral, c = contralateral): (1) N1-nodes: ipsilateral intrapulmonary, peribronchial and hilar nodes (S_{10-14}); (2) N2-nodes: subcarinal nodes (S_7), ipsilateral mediastinal and para-aortic nodes ($S_{2-6i, 8-9i}$); (3) N3-nodes: contralateral mediastinal and para-aortic nodes ($S_{2-6c, 8-9c}$), and supraclavicular nodes (S_1). A group of lymph node stations was rated positive, if at least one lymph node from one of the stations was considered to be metastatic.

Statistical analysis

Overall accuracy was calculated for primary tumour detection, T-staging, N-staging, group-wise assessment of lymph nodes and UICC staging. Sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) were calculated for primary tumour detection and group-wise assessment of lymph nodes. All measures were calculated along with their corresponding exact binomial 95% confidence intervals. A retrospective size analysis was performed based on the histology reports on those lymph node groups that were rated false negative by at least one of the four observers. Inter-observer agreement was assessed by calculating percentages of actually observed agreement without correction for effects of chance along with the corresponding exact binomial 95% confidence intervals. Statistical significance of the differences between the staging results obtained by PET/CT and MRI was tested using McNemar's test. Image quality ratings were compared by Wilcoxon signed-rank test. A *P* value of less than 0.05 was considered statistically significant.

Results

Whole-body MRI with DWIBS as well as PET/CT provided diagnostic image quality in all cases. Overall image quality for PET/CT was rated very good in 14 cases (45%), and good in the remaining 17 cases (55%), which was significantly better ($P=0.01$) than for MRI (very good, 7 [23%]; good, 20 [65%]; fair, 4 [13%]). Most important artefacts seen in MRI were ghosting artefacts from cardiovascular pulsations that affected morphological assessment of mediastinal structures.

Table 1 Detection of primary NSCLC by MRI and FDG-PET/CT (average values from two readers). Values *in parentheses* indicate 95% confidence intervals

Primary tumour detection	MRI	PET/CT
Sensitivity	26/28=93% (84–100%)	27.5/28=98% (93–100%)
Specificity	1.5/3=50% (0–100%)	1/3=33% (0–86%)
Accuracy	27.5/31=89% (77–100%)	28.5/31=92% (82–100%)
PPV	26/27.5=95% (86–100%)	27.5/29.5=93% (84–100%)
NPV	1.5/3.5=42% (0–93%)	1/1.5=75% (6–100%)

Assessment of individual MRI sequences revealed no significant differences ($P>0.1$) between DWI (very good, 6 [19%]; good, 21 [68%]; fair, 4 [13%]), T1-weighted TSE (very good, 7 [23%]; good, 18 [58%]; fair, 6 [19%]) and T2-weighted STIR (very good, 8 [15%]; good, 15 [48%]; fair, 8 [15%]).

The results of MRI and PET/CT for the detection of primary tumour lesions are displayed in Table 1. The individual numbers represent average values from two readers. Table 2 displays the T-staging results of the two techniques. Overall T-staging accuracy was higher for MRI (63%) than for PET/CT (56%). This difference, however, was not statistically significant ($P=0.6$).

Table 3 displays the N-staging results for MRI and PET/CT. N-staging accuracy was 66% for MRI and 71% for PET/CT. The results of group-wise assessment of metastatic lymph node involvement by MRI and PET/CT are shown in Table 4. A mean number of nine lymph node groups in total were rated false negative by each technique. In three cases, single metastases with sizes between 4 mm and 7 mm (maximum diameters measured by the pathologist on the surgical specimen) were missed by all four observers. In two cases, metastases with sizes of 8 mm and 15 mm were missed by both PET/CT readers, but were detected by one of the MRI readers. In another two cases, metastases with sizes of 7 mm and 15 mm that were detected by one of the PET/CT readers were

missed by both observers with MRI. Metastatic involvement of the remaining six lymph node groups was detected by at least one observer from each technique.

UICC staging results are displayed in Table 5. Comparison of methods by McNemar's test revealed no statistically significant difference between MRI and PET/CT for any of the calculated measures. The P value obtained for comparison of accuracy were 0.7 for primary tumour detection, 0.6 for T-staging, 0.65 for N-staging, 0.42 for assessment of individual lymph node groups and 0.4 for UICC staging. Figures 1 and 2 display two sample cases with discrepant diagnoses for T- and N-stage between the two modalities. A schematic overview on tumour detection and staging accuracies for both techniques is given in Fig. 3.

Observer agreement rates were 52% (34–70%) for T-staging, 68% (52–84%) for N-staging and 74% (59–89%) for UICC staging with MRI compared with 65% (48–82%), 68% (52–84%) and 90% (79–100%) with PET/CT. Inter-observer agreement generally tended to be higher for PET/CT than for MRI. However, differences were not statistically significant ($P=0.22$, 0.87 and 0.09).

Discussion

The objective of this study was to assess the diagnostic value of whole-body MRI with DWIBS for comprehensive

Table 2 T-staging results of MRI and FDG-PET/CT (average values from two readers). Resulting T-staging accuracy is 63% (95% CI, 46–80%) for MRI and 56% (39–73%) for FDG-PET/CT

T-staging		Histology				
		T0	T1	T2	T3	T4
MRI	total	3	10	14	2	2
	T0	1.5	2	0	0	0
	T1	1.5	6	3	0	0.5
	T2	0	1	10.5	1	0.5
	T3	0	1	0	1	0.5
PET/CT	T4	0	0	0.5	0	0.5
	T0	1	0.5	0	0	0
	T1	2	7.5	3	0	0.5
	T2	0	2	7.5	0.5	0.5
	T3	0	0	2.5	1	0.5
	T4	0	0	1	0.5	0.5

Table 3 N-staging results for MRI and FDG-PET/CT (average values from two readers). Resulting N-staging accuracy is 66% (95% CI: 49–83%) for MRI and 71% (55–87%) for FDG-PET/CT

N-staging		Histology			
		N0	N1	N2	N3
MRI	total	20	5	6	0
	N0	16.5	3	3	0
	N1	2	1.5	0.5	0
	N2	1	0.5	2.5	0
	N3	0.5	0	0	0
PET/CT	N0	17.5	2	3	0
	N1	2.5	2.5	1	0
	N2	0	0.5	2	0
	N3	0	0	0	0

Table 4 Group-wise assessment of metastatic lymph node involvement by MRI and FDG-PET/CT (average values from two readers). Values *in parentheses* indicate 95% confidence intervals

Lymph node assessment	N1 nodes		N2 nodes		Total (N1-N3)	
	MRI	PET/CT	MRI	PET/CT	MRI	PET/CT
Sensitivity [%]	45 (14-76)	55 (24-88)	42 (2-81)	33 (0-71)	44 (19-68)	47 (22-71)
Specificity [%]	86 (71-100)	88 (74-100)	92 (81-100)	98 (93-100)	93 (87-99)	96 (92-100)
Accuracy [%]	73 (57-88)	78 (63-92)	82 (68-96)	86 (73-98)	85 (72-97)	88 (76-99)
PPV [%]	65 (31-99)	66 (34-99)	58 (12-100)	83 (37-100)	61 (34-88)	71 (44-99)
NPV [%]	77 (59-94)	82 (66-97)	87 (74-100)	86 (73-99)	89 (82-96)	90 (83-96)

preoperative assessment of NSCLC in comparison with PET/CT with histopathology and cytology as the reference standards.

With respect to primary tumour detection, MRI and PET/CT both provided excellent results with sensitivity, accuracy and PPV of more than 89% for both techniques. The values obtained for sensitivity and accuracy are similar to previously reported results [9–12] which are 70–100% sensitivity and 72–100% accuracy for MRI, and 76–100% sensitivity and 74–100% accuracy for PET/CT, respectively. The values for specificity and NPV derived from the current study should be interpreted with caution because of the small sample size with only three non-malignant lesions. Specificity values for primary tumour detection reported in the literature are 96–97% for MRI and 79–82% for PET/CT [9–12].

Accuracy of T-staging is found to be slightly higher for MRI (63%) compared with 56% for PET/CT. However, this difference is far from being statistically significant ($P=0.6$). To the best of our knowledge, there are no published data on T-staging accuracy of MRI with diffusion weighting in NSCLC patients. The observed accuracy of 56% for PET/CT is in good accordance with the results from a recently published retrospective study [28]. In general, published values for T-staging accuracy of PET/CT cover a wide range

from 88% [3] down to only 39% in a study focused on stage IIIA disease [29]. In our study, most T-staging errors occurred with both techniques by mixing stages T1 and T2, which is of limited clinical relevance.

N-staging is a key issue in preoperative work-up of NSCLC patients. Until submission of this work, four studies had compared the value of diffusion-weighted MRI and PET/CT for lymph node assessment in NSCLC patients [11–14]. The published results from these studies were 67–91% sensitivity, 87–99% specificity and 80–98% accuracy for MRI, compared with 48–98% sensitivity, 89–97% specificity and 80–97% accuracy for PET/CT. The specificity and accuracy values observed in our study are in very good accordance with these results. However, sensitivity values calculated from our data are lower than the reported values for both techniques. It is well known that FDG-PET is considerably limited in detecting small lymph node metastases. A study that evaluated size dependence of PET/CT in N-staging of NSCLC patients reported a sensitivity of only 32% in lymph nodes <10 mm compared with 85% in lymph nodes >10 mm leading to a moderate sensitivity of 54% overall [30]. Another recently published study on PET/CT in early stage NSCLC also revealed low sensitivity for lymph node involvement of 44% [31]. In our study, approximately half of the false-negative lymph nodes were metastases smaller than 10 mm. In our opinion, this explains sufficiently the sensitivity results provided by PET/CT.

The question of spatial resolution limitations in NSCLC lymph node staging with diffusion-weighted MRI is discussed controversially in literature. The authors of one recently published study [11] claim that MRI may be superior to PET/CT in detecting small lymph node metastases, thus resulting in higher overall sensitivity and accuracy for MRI. Others [12, 13] see no relevant advantage of MRI in this respect, but describe higher accuracy and PPV of MRI compared with PET/CT owing to fewer false-positive findings. Sensitivity and PPV for metastatic lymph node involvement cannot be assessed with sufficient accuracy by our study because of the small sample size. Differences in accuracy, however, that are seen in our data between diffusion-weighted MRI and PET/CT regarding assessment of lymph node

Table 5 UICC staging results for MRI and FDG-PET/CT (average values from two readers). Resulting UICC staging accuracy is 66% (95% CI, 49–83%) for MRI and 74% (59–89%) for FDG-PET/CT

UICC-Staging		Histology			
		0	I-II	IIIA	IIIB-IV
MRI	total	3	17	6	5
	0	1.5	2	0	0
	I-II	1.5	12.5	3	1
	IIIA	0	1	3	0.5
PET/CT	IIIB-IV	0	1.5	0	3.5
	0	1	0.5	0	0
	I-II	2	14.5	2.5	0.5
	IIIA	0	1.5	3.5	0.5
	IIIB-IV	0	0.5	0	4

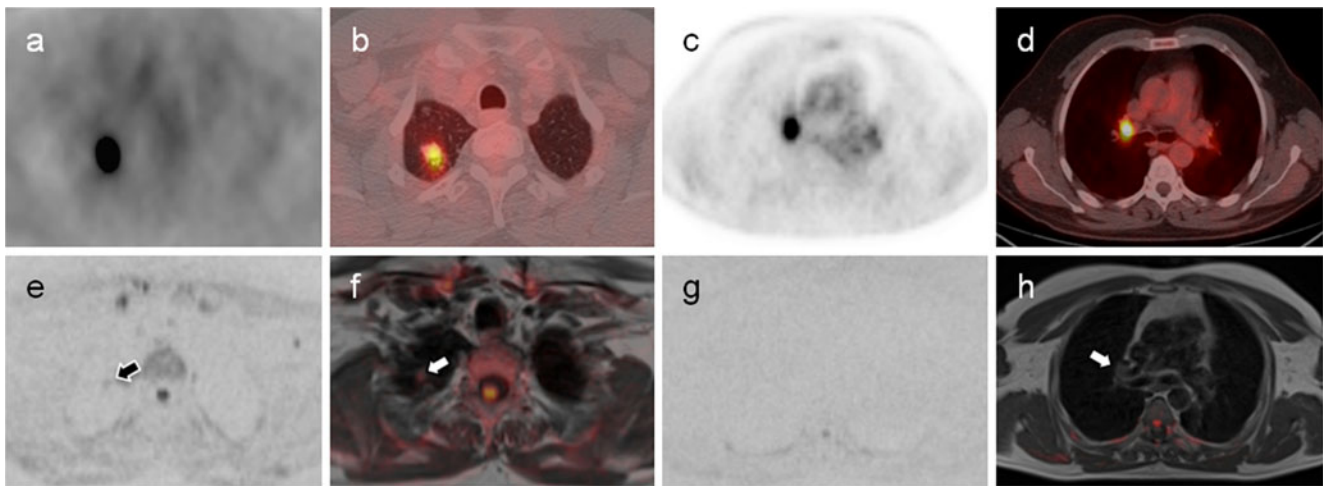


Fig. 1 Sample case of a 50-year-old man with primary tuberculosis. *Upper row:* FDG-PET source images in inverted grey scale (a, c), and combined PET/CT images after fusion of PET data on CT images in lung (b) and soft tissue window (d). *Lower row:* High b-value diffusion-weighted MRI source images in inverted grey scale (e, g), and combined images from high b-value diffusion-weighted and T1-weighted MRI data (f, h). This case was rated false positive for

malignancy (T1a N1, UICC IIA) by both PET/CT readers due to high FDG-uptake in both the primary lesion (a, b) and in a right hilar lymph node (c, d). On MRI, one of the readers considered the primary lesion to be malignant (T1) due to a small area of restricted diffusion within the primary lesion (arrows in e, f), whereas the second reader correctly identified the lesion as benign. Both MRI readers correctly described no suspicious lymphadenopathy (g, h)

groups and N-staging are far from being statistically significant ($P=0.42$ and 0.65). Thus, even considering the fact that the sample size included by Usuda et al. [11] is twice as high as in our study, we are not able to reproduce the very positive results for MRI given by this previous study.

Accuracy of UICC staging in NSCLC by means of diffusion-weighted MRI has to our knowledge only been investigated by one earlier study [11]. The reported values of 71% for DWI and 65% for PET/CT are in good accordance

with the results from our study (66% for DWI and 74% for PET/CT). The slight difference that is observed in favour of PET/CT is not reflected by the individual T- and N-staging results and is statistically not significant. The level of confidence ($P=0.4$) is almost equal to the number given by Usuda et al. ($P=0.44$) [11].

The method of diffusion-weighted whole-body MRI that has been evaluated in this study is from a technical perspective comparable to the methods that have been used by

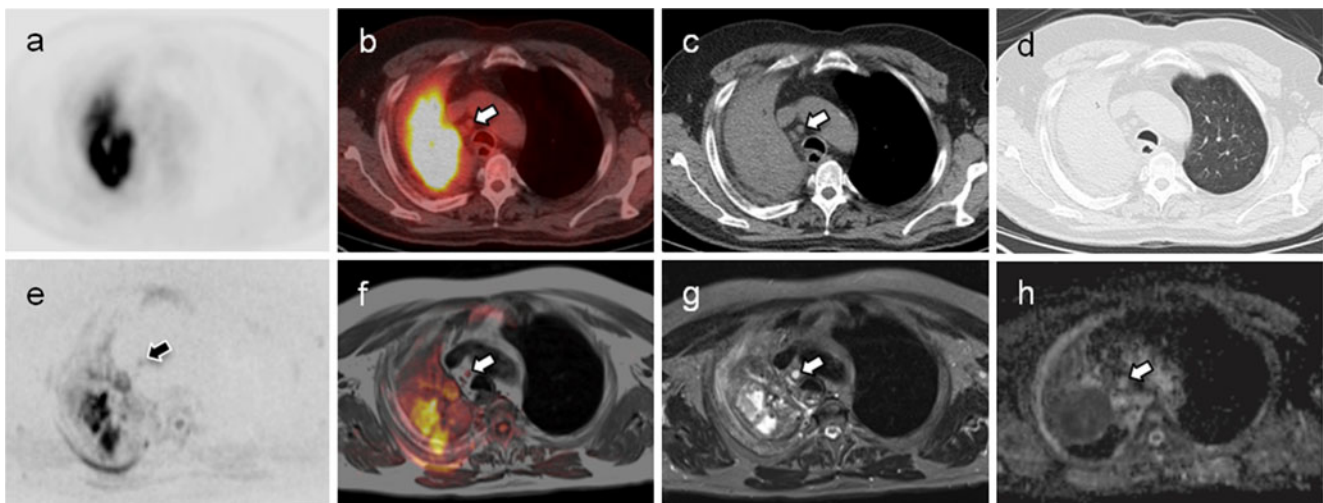


Fig. 2 Sample case of a 55-year-old man with pT2b pN1, UICC IIB squamous cell carcinoma. *Upper row:* FDG-PET source image (a), original CT images in soft tissue (c) and lung window (d), and combined PET/CT image (b). *Lower row:* High b-value diffusion-weighted source image (e), combined image from high b-value DWI and T1-weighted MRI (f), T2-weighted STIR image (g), and ADC-map (h). This case was rated false positive for N2-disease by one of the MRI readers due to increased signal

intensity in a right paratracheal lymph node (arrows) on both T2-weighted STIR and high b-value DWI. Increased ADC_{mean} ($1.9 \times 10^{-3} \text{ mm}^2/\text{s}$ measured with a 2D ROI in h) was interpreted by this reader as necrotic changes. PET/CT showed no increased FDG-uptake in this location (a/b). T-stage, however, was overestimated in the same patient by both PET/CT readers (T4 and T3) due to inflammatory changes in the surrounding lung parenchyma, but correctly assessed by both MRI readers

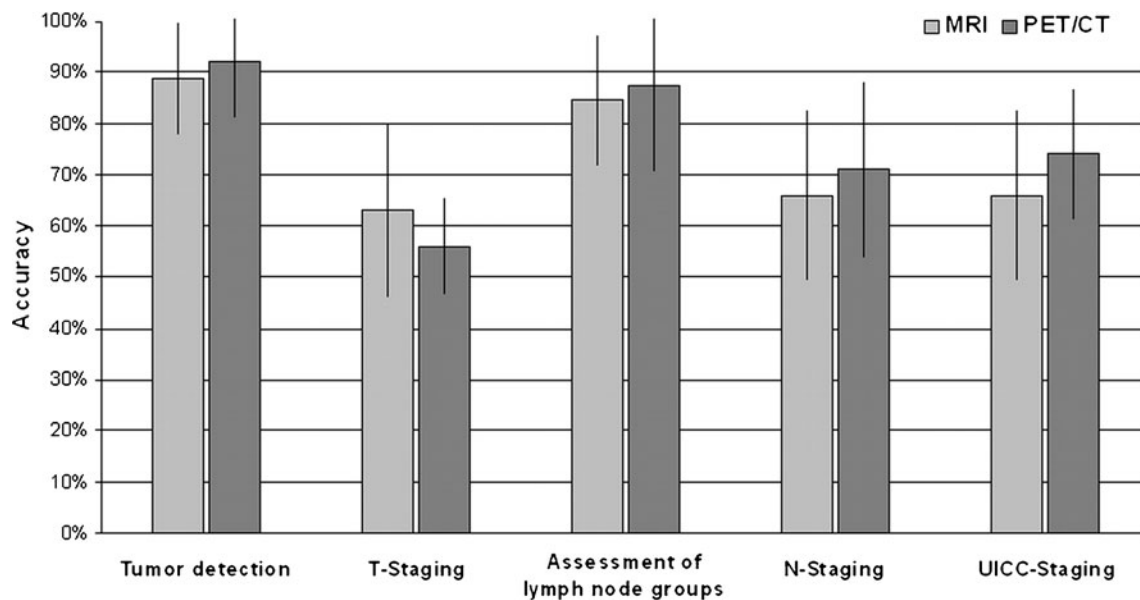


Fig. 3 Schematic overview of tumour detection and staging accuracies for MRI and PET/CT. The *error bars* represent 95% confidence intervals

previous authors. We chose a very simple approach for our protocol including two b values, free-breathing acquisition, STIR–fat suppression and a moderate number of averages to facilitate efficient whole-body coverage. Potential technical improvements as applied by other authors would be the use of more or higher b values, more elaborate methods of fat saturation such as spectral pre-saturation inversion recovery (SPAIR), and a higher number of averages. With reference to former studies, as cited above, we would expect most of these modifications to have only minor effects on overall staging accuracy. Artefacts from breathing motion and cardiovascular pulsations, however, were significantly more important in all MRI sequences applied in our study compared with PET/CT and represent a major issue in this context. The use of breath-holding, respiratory gating and ECG gating techniques may improve delineation of anatomical structures particularly in the central part of the chest and thus may further increase sensitivity for small central pulmonary lesions and mediastinal lymph nodes.

This study has several limitations. First to mention is the small patient number that particularly compromises the calculation accuracy of specificity and NPV for primary tumour detection and sensitivity and PPV for detection of lymph node metastases. Second, no valuable data on N3 lymph nodes could be obtained, as only patients with a potentially resectable stage of disease were included in the study. For the same reason, M-stage could not be assessed reliably as only two patients with solitary brain or lung metastases and curative intention for surgery were included. Thus, also the clinically important question of differentiating between UICC stages IA–IIIA and stages IIIB–IV has not been addressed. Another limitation of the study design is

the delay between PET/CT and surgery being systematically longer than between MRI and surgery. This might be interpreted as a bias in favour of MRI. However, we also observed two cases where the interpretation of MRI was impaired considerably by post-stenotic pneumonia or atelectasis that was not present at the time of the PET/CT examination. A systematic disadvantage for MRI that has to be discussed is reading expertise: radiologists and nuclear medicine physicians who acted as readers in this study had comparable expertise in their fields. However, compared with PET/CT, reading of DWIBS images is not yet standardised well and individual experience of readers with this new technique is still limited even for otherwise experienced radiologists. This is expressed by lower observer agreement for MRI compared with PET/CT in our study. Finally, the study outcome for both imaging techniques is limited by the fact that all staging results are based on the impression of a single board-certified radiologist or nuclear medicine physician. However, this reflects a quite common situation in clinical practise. It can reasonably be assumed that additional consensus readings in cases of discrepancy would likely have improved staging accuracies for both techniques.

In conclusion, this study has shown, in agreement with previously published studies, that whole-body MRI with DWIBS is a powerful method for staging of NSCLC and provides results comparable to the reference standard PET/CT. Thus, whole-body MRI with DWIBS may qualify as first line technique for staging of NSCLC when PET/CT is not available. However, as opposed to other authors, we are not yet convinced that there is clear evidence of a superiority of DWI with respect to lymph node assessment. We agree that the method of diffusion-weighted MRI has two intrinsic

technical advantages over FDG-PET, which are spatial resolution and soft tissue contrast. However, this potential has not been exploited to its full extent by today's routinely available applications. There is certainly a need for further technical improvement of both diffusion-weighted and conventional MRI sequences for optimised morphological and functional assessment of pulmonary and mediastinal structures.

Acknowledgements Preliminary results of this work and parts of the material enclosed have been presented at ECR 2012 as EPOS no. 4246.M. Klarhöfer is an employee of Siemens Switzerland Ltd., Healthcare Sector, Zürich, Switzerland. The authors acknowledge financial support from Guerbet Switzerland. The study sponsor played no role in matters of design, collection, analysis, interpretation of data and writing the report.

References

- Siegel R, Ward E, Brawley O, Jemal A (2011) Cancer statistics, 2011: the impact of eliminating socioeconomic and racial disparities on premature cancer deaths. *CA Cancer J Clin* 61:212–236
- Tolozan EM, Harpole L, McCrory DC (2003) Noninvasive staging of non-small cell lung cancer: a review of the current evidence. *Chest* 123:137S–146S
- Lardinois D, Weder W, Hany TF et al (2003) Staging of non-small-cell lung cancer with integrated positron-emission tomography and computed tomography. *N Engl J Med* 348:2500–2507
- Aquino SL, Kuester LB, Muse VV, Halpern EF, Fischman AJ (2006) Accuracy of transmission CT and FDG-PET in the detection of small pulmonary nodules with integrated PET/CT. *Eur J Nucl Med Mol Imaging* 33:692–696
- Roberts PF, Follette DM, von Haag D et al (2000) Factors associated with false-positive staging of lung cancer by positron emission tomography. *Ann Thorac Surg* 70:1154–1160
- Takahara T, Imai Y, Yamashita T, Yasuda S, Nasu S, Van Cauteren M (2004) Diffusion weighted whole body imaging with background body signal suppression (DWIBS): technical improvement using free breathing, STIR and high resolution 3D display. *Radiat Med* 22:275–282
- Matsumoto Y, Kuroda M, Matsuya R et al (2009) In vitro experimental study of the relationship between the apparent diffusion coefficient and changes in cellularity and cell morphology. *Oncol Rep* 22:641–648
- Kwee TC, Takahara T, Ochiai R, Nieuwelstein RAJ, Luijten PR (2008) Diffusion-weighted whole-body imaging with background body signal suppression (DWIBS): features and potential applications in oncology. *Eur Radiol* 18:1937–1952
- Mori T, Nomori H, Ikeda K et al (2008) Diffusion-weighted magnetic resonance imaging for diagnosing malignant pulmonary nodules/masses: comparison with positron emission tomography. *J Thorac Oncol* 3:358–364
- Ohba Y, Nomori H, Mori T et al (2009) Is diffusion-weighted magnetic resonance imaging superior to positron emission tomography with fludeoxyglucose F 18 in imaging non-small cell lung cancer? *J Thorac Cardiovasc Surg* 138:439–445
- Usuda K, Zhao XT, Sagawa M et al (2011) Diffusion-weighted imaging is superior to positron emission tomography in the detection and nodal assessment of lung cancers. *Ann Thorac Surg* 91:1689–1695
- Chen W, Jian W, H-tao L et al (2010) Whole-body diffusion-weighted imaging vs. FDG-PET for the detection of non-small-cell lung cancer. How do they measure up? *Magn Reson Imaging* 28:613–620
- Nomori H, Mori T, Ikeda K et al (2008) 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* 135:816–822
- Ohno Y, Koyama H, Yoshikawa T et al (2011) 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. *Radiology* 261:605–615
- Ohno Y, Koyama H, Onishi Y et al (2008) Non-small cell lung cancer: whole-body MR examination for M-stage assessment—utility for whole-body diffusion-weighted imaging compared with integrated FDG PET/CT. *Radiology* 248:643–654
- Satoh S, Kitazume Y, Ohdama S, Kimura Y, Taura S, Endo Y (2008) Can malignant and benign pulmonary nodules be differentiated with diffusion-weighted MRI? *AJR Am J Roentgenol* 191:464–470
- Uto T, Takehara Y, Nakamura Y et al (2009) Higher sensitivity and specificity for diffusion-weighted imaging of malignant lung lesions without apparent diffusion coefficient quantification. *Radiology* 252:247–254
- Takenaka D, Ohno Y, Matsumoto K et al (2009) Detection of bone metastases in non-small cell lung cancer patients: comparison of whole-body diffusion-weighted imaging (DWI), whole-body MR imaging without and with DWI, whole-body FDG-PET/CT, and bone scintigraphy. *J Magn Reson Imaging* 30:298–308
- Nakayama J, Miyasaka K, Omatsu T et al (2010) 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* 34:1–8
- Liu H, Liu Y, Yu T, Ye N (2012) Usefulness of diffusion-weighted MR imaging in the evaluation of pulmonary lesions. *Eur Radiol* 20:807–815
- Pauls S, Schmidt SA, Juchems MS et al (2010) Diffusion-weighted MR imaging in comparison to integrated [(18)F]-FDG PET/CT for N-staging in patients with lung cancer. *Eur J Radiol* 81:178–182
- Hasegawa I, Boiselle PM, Kuwabara K, Sawafuji M, Sugiura H (2008) Mediastinal lymph nodes in patients with non-small cell lung cancer: preliminary experience with diffusion-weighted MR imaging. *J Thorac Imaging* 23:157–161
- Matoba M, Tonami H, Kondou T et al (2007) Lung carcinoma: diffusion-weighted MR imaging—preliminary evaluation with apparent diffusion coefficient. *Radiology* 243:570–577
- Koşucu P, Tekinbaş C, Erol M et al (2009) Mediastinal lymph nodes: assessment with diffusion-weighted MR imaging. *J Magn Reson Imaging* 30:292–297
- Travis WD, Colby TV, Corrin B et al (1999) *Histological typing of lung and pleural tumours*, 3rd edn. Springer, Berlin Heidelberg
- UyBico SJ, Wu CC, Suh RD, Le NH, Brown K, Krishnam MS (2010) Lung cancer staging essentials: the new TNM staging system and potential imaging pitfalls. *Radiographics* 30:1163–1181
- Rusch VW, Asamura H, Watanabe H et al (2009) The IASLC lung cancer staging project: a proposal for a new international lymph node map in the forthcoming seventh edition of the TNM classification for lung cancer. *J Thorac Oncol* 4:568–577
- Subedi N, Scarsbrook A, Darby M, Korde K, Mc Shane P, Muers MF (2009) The clinical impact of integrated FDG PET-CT on management decisions in patients with lung cancer. *Lung Cancer* 64:301–307

29. Muehling B, Wehrmann C, Oberhuber A, Schelzig H, Barth T, Orend KH (2012) Comparison of Clinical and Surgical-Pathological Staging in IIIA Non-Small Cell Lung Cancer Patients. *Ann Surg Oncol* 19:89–93
30. Billé A, Pelosi E, Skanjeti A et al (2009) Preoperative intrathoracic lymph node staging in patients with non-small-cell lung cancer: accuracy of integrated positron emission tomography and computed tomography. *Eur J Cardiothorac Surg* 36:440–445
31. Li X, Zhang H, Xing L et al (2011) Mediastinal lymph nodes staging by (18)F-FDG PET/CT for early stage non-small cell lung cancer: A multicenter study. *Radiother Oncol* 102:246–250