Integrated FDG-PET/CT does not make invasive staging of the intrathoracic lymph nodes in non-small cell lung cancer redundant: a prospective study

K G Tournoy, S Maddens, R Gosselin, G Van Maele, J P van Meerbeeck, A Kelles

Background: Staging of non-small cell lung cancer (NSCLC) is important for determining choice of treatment and prognosis. The accuracy of FDG-PET scans for staging of lymph nodes is too low to replace invasive nodal staging. It is unknown whether the accuracy of integrated FDG-PET/CT scanning makes invasive staging redundant.

Methods: In a prospective study, the mediastinal and/or hilar lymph nodes in patients with proven NSCLC were investigated with integrated FDG-PET/CT scanning. Pathological confirmation of all suspect lymph nodes was obtained to calculate the accuracy of the fusion images. In addition, the use of the standardised uptake value (SUV) in the staging of intrathoracic lymph nodes was analysed.

Results: 105 intrathoracic lymph node stations from 52 patients with NSCLC were characterised. The prevalence of malignancy in the lymph nodes was 36%. The sensitivity of the integrated FDG-PET/CT scan to detect malignant lymph nodes was 84% and its specificity was 85% (positive likelihood ratio 5.64, negative likelihood ratio 0.19). SUV\text{max,} SUV\text{mean} and the SUV\text{max}/SUV\text{liver} ratio were all significantly higher in malignant than in benign lymph nodes. The area under the receiver operating curve did not differ between these three quantitative variables, but the highest accuracy was found with the SUV\text{max}/SUV\text{liver} ratio. At a cut-off value of 1.5 for the SUV\text{max}/SUV\text{liver} ratio, the sensitivity and specificity to detect malignant lymph node invasion were 82% and 93%, respectively.

Conclusion: The accuracy of integrated FDG-PET/CT scanning is too low to replace invasive intrathoracic lymph node staging in patients with NSCLC. The visual interpretation of the fusion images of the integrated FDG-PET/CT scan can be replaced by the quantitative variable SUV\text{max}/SUV\text{liver} without loss of accuracy for intrathoracic lymph node staging.

Staging non-small lung cancer (NSCLC) is an important part of the diagnostic course in patients with lung cancer since it guides the treatment modalities and predicts survival. While staging with computed tomography (CT) has an important role in the initial staging by providing excellent anatomical information on the extent of the primary tumour (T denominator), the CT scan has limited ability to differentiate between benign and malignant lymph nodes (N denominator). Whole body positron emission tomography (PET) with 18-fluoro-2-deoxy-D-glucose (FDG) has a higher accuracy for detecting intrathoracic lymph node metastasis, and demonstrates occult distant metastasis in approximately 10% of patients. According to other reports, CT and FDG-PET perform similarly in mediastinal staging. Because of an unacceptable rate of false positive and false negative findings, FDG-PET has been shown to be insufficiently accurate to replace invasive lymph node staging on tissue specimens. In addition, the visual localisation of intrathoracic lymph nodes with FDG-PET is not always unequivocal because of the low spatial resolution of the PET images.

Integrated FDG-PET/CT scans theoretically overcome this problem because of the co-acquisition of CT and FDG-PET images resulting in so-called fusion images. Lardinois et al. showed that, in 50 patients with NSCLC, the integrated FDG-PET/CT scan was more accurate than FDG-PET alone for nodal staging; however, no difference in accuracy was noted when integrated FDG-PET/CT scans were compared with CT scans alone. Several investigators have indicated that the maximum standardised uptake value (SUV\text{max}) of the primary tumour is a prognostic factor in patients with NSCLC. With the integrated FDG-PET/CT scan, the application of this quantitative technique at the level of the lymph node becomes more precise since the anatomical borders of the lymph nodes can be exactly identified for determining the region of FDG uptake. From a clinical point of view, an accuracy or a negative/positive predictive value of at least 90–95% for the integrated FDG-PET/CT scan is required to make invasive staging redundant. To evaluate whether tissue-confirmed lymph node staging by surgery or echo-endoscopy can be avoided, we prospectively assessed the accuracy of the integrated FDG-PET/CT scan in the nodal staging of NSCLC. In addition, we investigated whether an objective measure of FDG uptake based on SUV values could substitute for the subjective interpretation of the fusion images.

METHODS

Patients

Consecutive patients with suspected or pathologically proven primary NSCLC were eligible if a tissue specimen from at least one of the intrathoracic lymph nodes was available and if they underwent an integrated FDG-PET/CT scan. All investigations were done before the start of any treatment, and both the integrated PET/CT scan and examination of lymph node tissue were done before the start of any treatment, and both the integrated PET/CT scan and examination of lymph node tissue

Abbreviations: CT, computed tomography; EBUS-TBNA, endobronchial ultrasound with real-time guided transbronchial needle aspiration; EUS-FNA, endoscopic ultrasound with real-time guided fine needle aspiration; FDG, 18-fluoro-2-deoxy-D-glucose; NSCLC, non-small cell lung cancer; PET, positron emission tomography; ROC, receiver operating curve; SUV, standardised uptake value
were performed within 14 days of each other. The study was approved by the ethics committee of Ghent University Hospital.

**Pathology of primary lung tumours and of intrathoracic lymph nodes**

A tissue specimen of the primary tumour was obtained for pathological examination by either bronchoscopy, CT-guided transbronchial puncture or a surgical procedure (thoracotomy or video-assisted thoracoscopic surgery). For intrathoracic lymph nodes, a tissue sample was obtained either by mediastinoscopy, surgical resection or by linear endoscopic ultrasound. The latter consisted of either oesophageal endoscopic ultrasound with real-time guided fine needle aspiration, endobronchial endoscopic ultrasound with real-time guided tranbronchial needle aspiration.

**Results**

The study included 52 patients with NSCLC whose characteristics are shown in table 1. Patients were recruited between November 2005 and April 2006. A diagnosis of NSCLC was obtained in the primary tumour specimen in 67% of cases. For the others, the diagnosis was extrapolated from the malignant cells found in the lymph node specimen. 52% of the patients had at least one malignant intrathoracic lymph node. Integrated FDG-PET/CT analysis and tissue specimens were obtained in the primary tumour specimen in 67% of cases. For the others, the diagnosis was extrapolated from the malignant cells found in the lymph node specimen. 52% of the patients had at least one malignant intrathoracic lymph node. Integrated FDG-PET/CT analysis and tissue specimens were obtained in the primary tumour specimen in 67% of cases. For the others, the diagnosis was extrapolated from the malignant cells found in the lymph node specimen. 52% of the patients had at least one malignant intrathoracic lymph node. Integrated FDG-PET/CT analysis and tissue specimens were obtained in the primary tumour specimen in 67% of cases. For the others, the diagnosis was extrapolated from the malignant cells found in the lymph node specimen. 52% of the patients had at least one malignant intrathoracic lymph node.

**Statistical analysis**

A clinical research form was completed for each patient and integrated in an electronic database for analysis with SPSS 14.0 (SPSS Inc, Chicago, Illinois, USA). Except for the demographic data, calculations were primarily performed at the single lymph node level. The test characteristics of the CT scan and integrated FDG-PET/CT scan were also calculated at the individual patient level. Values are expressed as mean (SD). Continuous variables were compared with the Mann-Whitney U test and Kruskal-Wallis test. Comparison of proportions (accuracy analysis) was done with a Z-test. A two-sided p value of <0.05 was considered statistically significant. The receiver operating curves (ROC) and the ROC areas, the latter being a parameter to measure how well a certain variable can distinguish between benign and malignant lymph nodes, were analysed with SPSS 14.0 and compared with Medcalc Version 9.1.0.1.

**Results**

The study included 52 patients with NSCLC whose characteristics are shown in table 1. Patients were recruited between November 2005 and April 2006. A diagnosis of NSCLC was obtained in the primary tumour specimen in 67% of cases. For the others, the diagnosis was extrapolated from the malignant cells found in the lymph node specimen. 52% of the patients had at least one malignant intrathoracic lymph node. Integrated FDG-PET/CT analysis and tissue specimens were obtained in the primary tumour specimen in 67% of cases. For the others, the diagnosis was extrapolated from the malignant cells found in the lymph node specimen. 52% of the patients had at least one malignant intrathoracic lymph node.
necessary because of a negative endoscopic ultrasound examination.

**Standardised uptake values for the primary tumour**

Table 2 shows the mean values of the different SUV variables for the primary tumours. There was no difference in SUV max, SUV mean or in the ratio SUV max/SUV liver between the different histological subtypes of NSCLC (Kruskal-Wallis test), nor between squamous and non-squamous carcinomas. Both standard deviation (table 2) and variance (data not shown) of the SUV variables indicate that the spread of the SUV max was considerably larger than for SUV mean or the SUV max/SUV liver ratio. In addition, lung tumours ≤3 cm had a statistically significant lower SUV max, SUV mean and SUV max/SUV liver than tumours >3 cm. We did not find a significant correlation between the four T stages and the respective SUV values (data not shown).

**Table 2** Mean (SD) PET/CT data for the primary tumour in patients with non-small cell lung cancer (NSCLC)

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>SUVmax (mm)</th>
<th>SUVmean (mm)</th>
<th>SUVmax/SUVliver</th>
</tr>
</thead>
<tbody>
<tr>
<td>All NSCLC</td>
<td>52</td>
<td>8.34 (4.02)</td>
<td>4.74 (2.57)</td>
<td>4.78 (2.18)</td>
</tr>
<tr>
<td>Adenocarcinoma</td>
<td>17  (33%)</td>
<td>7.50 (3.79)</td>
<td>4.72 (3.12)</td>
<td>4.06 (2.22)</td>
</tr>
<tr>
<td>Squamous cell carcinoma</td>
<td>20  (38%)</td>
<td>9.04 (4.61)</td>
<td>4.91 (2.85)</td>
<td>5.37 (2.30)</td>
</tr>
<tr>
<td>Large cell carcinoma</td>
<td>10  (19%)</td>
<td>8.32 (3.41)</td>
<td>4.46 (2.10)</td>
<td>4.88 (1.33)</td>
</tr>
<tr>
<td>Adenosquamous carcinoma</td>
<td>2  (10%)</td>
<td>8.47 (4.08)</td>
<td>4.71 (1.82)</td>
<td>5.64 (2.21)</td>
</tr>
<tr>
<td>T1 tumours (&lt;3 cm)</td>
<td>16  (31%)</td>
<td>6.65 (3.42)</td>
<td>3.71 (1.87)</td>
<td>3.77 (1.52)</td>
</tr>
<tr>
<td>T2–4 tumours (&gt;3 cm)</td>
<td>36  (69%)</td>
<td>9.10 (4.09)</td>
<td>5.20 (2.73)</td>
<td>5.23 (2.29)</td>
</tr>
</tbody>
</table>

Mann-Whitney U test p = 0.026 p = 0.036 p = 0.008

SUV, standardised uptake value.

**Accuracy of CT and integrated FDG-PET/CT for staging lymph nodes**

The accuracy of the CT findings and of the integrated FDG-PET/CT fusion images for staging intrathoracic lymph nodes is shown in table 3. The prevalence of malignant lymph nodes was 36%. Whereas the sensitivity of CT and integrated FDG-PET/CT scans to detect malignant lymph node invasion was equal (both 84%), integrated FDG-PET/CT scanning had a specificity of 85% compared with 61% for the CT scan. The accuracy for detecting malignant intrathoracic lymph nodes was 69% and 85%, respectively, for CT scans and integrated FDG-PET/CT scans. These figures do not differ significantly from the accuracy of the CT scan (74%) and integrated FDG-PET/CT scan (84%) calculated at the level of the individual patient. The negative predictive value of small lymph nodes (short axis ≤10 mm) without uptake of FDG was 91% (95% CI 77% to 97%; n = 43); the positive predictive value of enlarged lymph nodes (short axis >10 mm) with FDG uptake was 79% (95% CI 62% to 90%; n = 38); the negative predictive value of enlarged lymph nodes without uptake of FDG was 90% (95% CI 60% to 98%; n = 20); and the positive predictive value of small lymph nodes with FDG uptake was 50% (95% CI 6% to 93%; n = 4).

**Table 3** Accuracy of interpretation of computed tomography (CT) scans versus integrated whole body positron emission tomography with 18-fluoro-2-deoxy-D-glucose/CT (FDG-PET/CT) fusion images for staging of intrathoracic lymph nodes in patients with non-small cell lung cancer (NSCLC)

<table>
<thead>
<tr>
<th></th>
<th>All lymph nodes n = 105</th>
<th>Benign lymph nodes n = 67 (64%)</th>
<th>Malignant lymph nodes n = 38 (36%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CT read-out</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CT positive</td>
<td>58 (55%) 26 (39%) 32 (84%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CT negative</td>
<td>47 (45%) 41 (61%) 6 (16%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CT specificity</td>
<td>84 (68-93)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CT LR+</td>
<td>61 (48-73) 2.17 (1.56–3.02)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CT LR−</td>
<td>0.26 (0.12–0.55)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PET/CT read-out</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PET/CT positive</td>
<td>42 (40%) 10 (15%) 32 (84%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PET/CT negative</td>
<td>63 (60%) 57 (85%) 6 (16%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PET/CT sensitivity</td>
<td>84 (68-93)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PET/CT specificity</td>
<td>85 (74-92)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PET/CT LR+</td>
<td>5.64 (3.13–10.16)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PET/CT LR−</td>
<td>0.19 (0.09–0.39)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

LR+, positive likelihood ratio; LR−, negative likelihood ratio.
Statistical analysis CT vs integrated PET/CT read-out: sensitivity, NS; specificity, p = 0.002; positive likelihood ratio, p<0.001; negative likelihood ratio, NS.
68% and a specificity of 93%. The SUVmax/SUVliver ratio had at a cut-off value of 1.5, the highest accuracy with a sensitivity of 82% and a specificity of 93%.

DISCUSSION

In this prospective study we have analysed the value of the integrated FDG-PET/CT scan for assessing intrathoracic lymph nodes in patients with NSCLC. Its accuracy for predicting malignant lymph node invasion was found to be higher than with a CT scan. We also found that the SUVmax/SUVliver ratio, a quantitative measure of FDG uptake, predicts malignant invasion in the intrathoracic lymph node with an accuracy comparable to that obtained with the integrated FDG-PET/CT fusion images. However, neither the interpretation of the integrated FDG-PET/CT fusion images nor the calculated cut-offs enabled us to continue using tissue-based lymph node staging in NSCLC because all the test characteristics were below the threshold of 95% at which malignant lymph node invasion can be confidently ruled in or out. Therefore, neither a positive nor a negative integrated FDG-PET/CT scan allows the clinician to confidently predict whether or not an intrathoracic lymph node is malignant.

The advantage of the integrated FDG-PET/CT scan over the CT scan as well as the FDG-PET scan in thoracic oncology is the fact that the spatial resolution of the integrated scan is much higher than with the FDG-PET scan. Because of co-acquisition and fusion, FDG uptake can be exactly localised even in relatively small focal abnormalities such as intrathoracic lymph nodes. This allows a detailed analysis of the value of SUV-based variables within a certain area of interest defined by the anatomical borders.

It has been shown in several retrospective series that the SUVmax obtained with FDG-PET scanning of the primary tumour predicts both stage and survival of the patient with NSCLC. A correlation between T stage and SUVmax was found by Cerfolio et al for T1 and T2 tumours but not for T3 and T4 tumours. The same authors also suggested that squamous cell carcinomas have a higher SUVmax than other pathological subtypes. SUVmax was proposed as a variable because it was suggested that its reproducibility is better than for SUVmean. However, others have suggested that SUVmean as well as the SUVmax/SUVliver ratio give more stable and reliable results.

In our prospective series we analysed these SUV-based characteristics of the primary tumour with integrated
FDG-PET/CT scanning and found that the clustering of SUV\textsubscript{mean} and the SUV\textsubscript{max}/SUV\textsubscript{liver} ratio was better than that of SUV\textsubscript{max}. This suggests that the former two might be more reliable and stable than SUV\textsubscript{max}. Although our data also confirm that T1 tumours had a lower SUV\textsubscript{max} than T2–4 tumours,\(^1\) we did not observe a difference between the different pathological subtypes.

For lymph node staging, we found that the integrated FDG-PET/CT scan performed better than the CT scan in terms of specificity (85% vs 61%, \(p = 0.002\)) but not sensitivity (84% for both). It is already known that FDG-PET scans are better than CT scans for staging of intrathoracic lymph nodes.\(^1\) However, Shim \textit{et al.}\(^6\) directly compared the accuracy of CT and integrated FDG-PET/CT scanning and suggested that the integrated FDG-PET/CT scan is superior for lymph node staging. Our data confirm this finding and further suggest that a visual interpretation of the fusion images is not sufficiently accurate to avoid invasive tissue specimen-based lymph node staging. Because the implication of a false positive staging evaluation is a missed opportunity for surgical care, we confirm that, as for FDG-PET, a positive integrated FDG-PET/CT result does not confidently rule in lymph node invasion. de Langen \textit{et al.}\(^5\) found that lymph nodes measuring 10–15 mm without FDG uptake had a post-test probability for malignant invasion of only 5%, so they suggested that these patients do not need further mediastinal investigation. Since the absolute number of false negative lymph nodes in our series was low, we are cautious about drawing firm conclusions on the negative predictive value of integrated FDG-PET/CT scans. Nevertheless, we calculated a specificity of 85% and a negative predictive value of 90%, which compares with the findings of Shim \textit{et al.}\(^6\)

We also determined whether the accuracy of intrathoracic lymph node staging could be improved by calculating SUV variables on the fusion images. As expected, we found that enlarged malignant lymph nodes have significantly higher SUV values than benign ones. With an FDG-PET scan it was proposed that a SUV\textsubscript{max} cut-off of 4.4 was the best threshold to discriminate between benign and malignant lymph nodes.\(^7\) With the integrated FDG-PET/CT scan, we calculated ROCs for SUV\textsubscript{mean} and also for SUV\textsubscript{max} and the SUV\textsubscript{max}/SUV\textsubscript{liver} ratio for lymph node staging and found that optimal cut-off points for all three variables could be identified. For SUV\textsubscript{max} we found that the optimal threshold for discriminating malignancy was 2.9. The discrepancy between published data\(^7\) and our findings could be partly attributed to the fact that FDG-PET techniques have evolved over the past 10 years resulting in more sensitive cameras, and also to the fact that different technology was used for the attenuation correction and for the calculation of the SUV data. For an FDG-PET examination, an external cesium source is used for calculating the attenuation correction while, for an FDG-PET/CT camera, the CT data are used to calculate the attenuation correction. This is the reason why we also included a ratio as variable (SUV\textsubscript{max}/SUV\textsubscript{liver}) so that camera-specific variables could be omitted. In addition, the lower SUV\textsubscript{max} threshold in our study could also be related to the fact that we were able to delineate exactly the nodal focus on the integrated FDG-PET/CT fusion images.

More importantly, our results indicate that a thoracic oncology clinic working with FDG-PET or with integrated FDG-PET/CT should not automatically rely on published SUV-based cut-off values but should determine their own machinespecific cut-off points. Although at a cut-off point of 1.5 the accuracy of the SUV\textsubscript{max}/SUV\textsubscript{liver} ratio was comparable to the one obtained with the fusion images, calculated SUV values also do not replace lymph node tissue diagnosis.

This study has a number of limitations. A number of methods encompassing echo-endoscopy and surgical staging procedures were used to obtain lymph node pathology. Although all negative echo-endoscopies were confirmed with surgery, it is clear that patients who had proven N2/N3 disease after echo-endoscopy did not have full pathological mapping of the mediastinum. In addition, although over 100 lymph nodes were characterised, it is inevitable that—for the assessment of test variables—false negative and false positive findings are calculated on relatively small numbers. As discussed above, this means that the interpretation of negative predictive values of a small lymph node which is FDG-PET/CT negative should be done with caution. Furthermore, the data were not corrected for possible intrasubject correlation between nodes. Although re-calculation of the test characteristics at the patient level did not show a difference from the data obtained at the single lymph node level, this must be taken into account when interpreting the data.

In conclusion, integrated FDG-PET/CT scanning has an overall accuracy which is too low to replace invasive intrathoracic lymph node staging in patients with NSCLC. The visual interpretation of the fusion images of the integrated FDG-PET/CT scan can be replaced by the quantitative variable SUV\textsubscript{max}/SUV\textsubscript{liver} without loss of accuracy for staging of intrathoracic lymph nodes.

ACKNOWLEDGEMENTS
This is an investigator initiated study.

Authors’ affiliations
K G Tournoy, S Maddens, J P van Meerbeeck, Department of Respiratory Medicine, Ghent University Hospital, Ghent, Belgium
R Gosselin, Department of Radiology, Ghent University Hospital, Ghent, Belgium
G Van Maele, Department of Medical Statistics, Ghent University Hospital, Ghent, Belgium
A Kelles, Department of Nuclear Medicine, Ghent University Hospital, Ghent, Belgium

Funding: None.

Competing interests: None.

REFERENCES
COPD and MI: a bad combination


In this observational prospective multicentre study that enrolled 2481 patients, the authors tried to highlight the relationship between chronic obstructive pulmonary disease (COPD) and a number of outcomes following an acute myocardial infarction (MI).

One-year mortality and rehospitalisation rates were significantly higher in patients with COPD than without (15.8% vs 5.7%) and (48.7% vs 38.6%), respectively. In addition, patients with COPD had worse health status at baseline as well as at 1 year and a trend toward a higher prevalence of angina at 1 year. Not surprisingly, patients with COPD had fewer coronary revascularisation procedures and less use of β blockers, aspirin and cardiac rehabilitation.

With the well-documented role of β blockers in improving survival post-MI even in patients with COPD, it is possible that more frequent usage of these agents, particularly the cardioselective agents, in the study population could have led to a better 1 year survival than observed and possibly a lower prevalence of angina, again at 1 year.

The study used clinical rather than pulmonary function test definition of COPD and as a consequence could have misclassified some patients with possible inclusion of non-COPD patients, for example those with asthma. Furthermore, the outcomes could have been influenced by the severity of COPD, which the authors were unable to classify. Nevertheless, the data suggest that patients with acute MI and COPD have a substantially worse prognosis in terms of mortality and health status. This cohort of patients therefore warrants careful attention and closer follow-up.

Nabil M Al Lawati
Clinical Fellow, University of British Columbia, Canada; drnabilm@yahoo.com