LUNG CANCER

Prospective use of serial questionnaires to evaluate the therapeutic efficacy of $^{18}$F-fluorodeoxyglucose (FDG) positron emission tomography (PET) in suspected lung cancer

G J Herder, H van Tinteren, E F Comans, O S Hoekstra, G J Teule, P E Postmus, U Joshi, E F Smit

Background: A study was undertaken to study the effect of $^{18}$F-fluorodeoxyglucose (FDG) positron emission tomography (PET) on the diagnosis and management of clinically problematic patients with suspected non-small cell lung cancer (NSCLC).

Methods: A prospective before-after study was performed in a cohort of all 164 patients (university/community settings) referred for PET between August 1997 and July 1999. PET was restricted to cases where non-invasive tests had failed to solve clinical problems. The impact on diagnostic understanding and management was assessed using questionnaires (intended treatment without PET, actual treatment choice after PET, post hoc clinical assessment).

Results: Diagnostic problems especially pertained to unclear radiological findings (n=112; 63%), mediastinal staging (n=36; 20%), and distant staging issues (n=16; 9%). PET findings were validated by reviewing medical records. PET had a positive influence on diagnostic understanding in 84%. Improved diagnostic understanding solely based on PET was reported in 26%. According to referring physicians, PET resulted in beneficial change of treatment in 50%. Cancelled surgery was the most frequent change in treatment after PET (35%).

Conclusion: FDG PET applied as “add on” technology in patients with these clinical problems appears to be a clinically useful tool, directly improving treatment choice in 25% of patients. The value of increased confidence induced by PET scanning requires further evaluation.

Medical imaging technology is rapidly expanding and the role of each modality is being redefined constantly. Positron emission tomography (PET) using $^{18}$F-fluorodeoxyglucose (FDG) has emerged as an accurate imaging modality in patients with lung cancer. Potential clinical indications include the differential diagnosis of benign versus malignant disease, initial (preoperative) staging, evaluation of suspected recurrences, and follow up after treatment. The use of PET in clinical practice is based predominantly on studies of technical performance and diagnostic accuracy. To ensure the appropriate use of PET, such studies should be followed by an analysis of the impact of PET on management decisions, outcomes of care, and cost effectiveness.

In the northwestern part of the Netherlands where this study was performed, a single PET scanner serves 2.7 million inhabitants with 50% of its time slots available for clinical purposes. To restrict the use of PET to those patients who may benefit most, a programme has been developed to evaluate the clinical usefulness of PET, investigating the cost effectiveness of performing PET on a routine basis in the preoperative staging of non-small cell lung cancer (NSCLC) and its impact as an “add on” technique in specific problem cases. To measure the clinical value of PET in the latter group, we performed a prospective before-after study in a cohort of clinically problematic cases, typically after an extensive conventional work-up. This study design was used during the early studies of computed tomographic (CT) scanning by Wittenberg et al, and allows a systematic assessment of the impact of a test on diagnostic understanding as well as on patient management within the clinical context.

METHODS
To be eligible for PET scanning, patients had to have suspected or proven NSCLC with a diagnostic problem which, according to the referring physician, could not be solved by conventional methods alone and in which the PET result might affect patient management. In an attempt to restrict PET scanning to such cases, referrals were only accepted after discussion of the case between the physician and the staff nuclear medicine physician in charge at the Clinical PET Center of the VU University Medical Center. PET scanning therefore typically followed an extensive conventional work-up. All patients routinely underwent laboratory tests, bronchoscopy, chest radiography and CT scanning extending from the neck to the upper abdomen (including liver and adrenal glands). Additional diagnostic tests were performed in cases with signs and symptoms suggestive of distant metastatic disease. Patients entered in randomised or response monitoring trials were not included in the present report.

Assessment of clinical value
The impact of PET on diagnostic understanding and treatment choice was investigated using three questionnaires (fig 1). These questionnaires were completed by the referring physician before PET scanning, shortly after PET scanning, and about 6 months after PET scanning, respectively. In the first questionnaire, information was requested regarding the histological diagnosis (if known), a definition of the current diagnostic problem, a differential diagnostic consideration, the results of diagnostic tests already performed, and any planned diagnostic tests. In addition, the referring physician...
concerned were asked to review the cases in question and to revise the overall clinical value rating accordingly and these data were used in the analysis. In the case of PET negative—that is, suspected benign—lesions, follow up was extended beyond 6 months by examining the medical records of these patients.

Management changes
Treatment (management) changes were considered “major” if treatment changed from one modality to another—for example, from medical to surgical/radiation/no treatment or vice versa—and “minor” if treatment changed within a modality—for example, altered medical, surgical or radiotherapy approach.

PET imaging
Whole body FDG PET scans were performed with a dedicated PET scanner (ECAT EXACT HR+, CTI/Siemens). Emission scans, typically extending from mid skull to mid femur, were performed in 2D mode, approximately 60 minutes after intravenous injection of 370 MBq (10 mCi) FDG. Patients were asked to fast for at least 6 hours before the PET study. Oral intake of water was encouraged.

PET scans were corrected for decay, scatter, and random data. Scans were reconstructed as 128 × 128 matrices using filtered back projection with a Hanning filter (cut off 0.5 cycles/pixel) resulting in a transaxial spatial resolution of 7 mm at full width half maximum. If possible, CT scan data were used for more precise anatomical localisation of PET abnormalities suspected as being malignant.

Referring physicians were informed by telephone of the result of the PET scan and advice to the next step. Clinicians were urged to verify clinically decisive PET findings by conventional means (histology, imaging, follow up) and to ignore unconfirmed hot spots. PET findings were retrospectively validated by examination of the medical records of the patients. Histopathological and clinical follow up findings that showed a benign or malignant course were considered as a valid reference test.

Statistical analysis
Differences in diagnostic understanding or treatment choice between the three indications were tested by means of a two sided Kruskal-Wallis test. The Wilcoxon-Mann-Whitney test was used to test differences between two samples. Changes in
treatment plans before and after PET were tested by the marginal homogeneity test.11

**RESULTS**

During a 23 month period 179 patients with suspected NSCLC were referred for PET scanning. The referring physicians included pulmonologists (76%), oncologists (7%), internists (6%), radiotherapists (6%), neurologists (3%), and surgeons (1%) from 21 different university and community hospitals. Questionnaires were returned from 178 (99%) and a fully completed set of questionnaires (all items answered) was obtained for 136 (76%) patients. Specifically, questionnaire 1 was fully completed for 83% of the patients, questionnaire 2 for 92%, and questionnaire 3 for 98% of patients. Indications for PET could be subdivided into six groups: unclear radiological abnormality (including solitary pulmonary nodules and lung masses, n=112; 63%), staging of the mediastinum (n=36; 20%), distant staging issues (n=16; 9%), response monitoring (n=5; 2.8%), suspected recurrence (n=5; 2.8%), and unknown primary (n=5; 2.8%). The present report focuses on the first three clinical indications.

In these 164 patients the clinical work-up before PET included laboratory tests, chest radiography, CT scan of the chest (including liver and adrenal glands), and bronchoscopy.11 In patients with distant staging problems (n=16) the work-up before PET consisted of bone scintigraphy and radiographic studies in the three patients with clinical concerns about skeletal metastases; CT evaluation of the abdomen typically preceded referrals with suspect adrenal enlargement or liver lesions in which biopsy specimens were not considered feasible or had been inconclusive. In two patients in whom a chest CT scan had shown additional and indeterminate pulmonary lesions, bronchoscopic examination had been negative and it was not considered feasible to take biopsy specimens. In five patients with potentially solitary brain metastases, dissemination tests had included CT scanning (brain, chest, liver and adrenal glands) and bone scintigraphy. In general, the work-up of patients with unclear radiological findings before PET scanning conformed to national guidelines.11

The diagnostic problems concerning mediastinal staging leading to referral for PET (instead of invasive mediastinal staging) included former mediastinoscopy, thoracotomy or radiotherapy, indeterminate invasive staging results, medical inoperability, and “to determine the most appropriate surgical approach”. After careful evaluation we were unable to identify a specific reason for choosing PET scanning as opposed to mediastinoscopy to determine mediastinal lymph node involvement in 10 patients.

In 29 of the 179 patients the initially formulated management plans (to be carried out if PET had not been available) were not consistent with the final assessment of the impact of PET. For example, the physician’s written plan before PET was to perform a thoracotomy, and a thoracotomy was indeed performed but treatment choice was rated as 5 (PET was very important compared with other factors leading to a beneficial change in treatment). Such inconsistent assessments were revised by the referring physicians (specifically with respect to questionnaire 3) and corrected in 28 cases.

<table>
<thead>
<tr>
<th>Table 1: Impact of PET on diagnostic understanding (DU ratings defined in box 1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DU=1</td>
</tr>
<tr>
<td>-----------------</td>
</tr>
<tr>
<td>Radiological abnormality</td>
</tr>
<tr>
<td>Mediastinal staging</td>
</tr>
<tr>
<td>Distant staging</td>
</tr>
<tr>
<td>Overall</td>
</tr>
</tbody>
</table>

**Diagnostic understanding**

The impact of PET on diagnostic understanding was analysed for each clinical indication (table 1). Overall, PET was solely responsible for improved diagnostic understanding (DU=5) in 26% (95% CI 19 to 33) of the patients and substantially contributed to diagnostic understanding (DU=4) in 58% (95% CI 50 to 65). The effect of the PET result on diagnostic understanding was confusing and led to additional tests (DU=1) in 3% (95% CI 1 to 6) and had no or little effect (DU=3) in 9% (95% CI 5 to 15). The impact of PET on diagnostic understanding was not significantly different for the three clinical indications (p=0.45). There was no significant difference (p=0.85) in diagnostic understanding ratings between PET scans indicating malignancy where the tumour was finally proved to be malignant (true positives) and scans indicating benign disease where the lesion proved to be benign (true negatives). To evaluate the presence of a potential clinical learning curve of incorporating PET scanning results, we compared the diagnostic understanding rating of “early” patients (the first five patients) referred by a particular physician to the ratings of later patients (the sixth and subsequent patients). The ratings in later patients tended to be significantly higher (p=0.0192).

**Diagnostic accuracy**

Of the patients referred to resolve unclear radiological findings, 76 had a positive PET scan result which proved to be true positive in 68 patients (89%). Thirty six patients had a negative scan reading—that is, no focally enhanced FDG uptake suspicious of malignancy—which proved to be correct (true negative) in 34 patients (94%) either by “wait and see” policy (n=32) or surgery (n=2). The mean duration of follow up in these patients was 20 months (range 6–36). In two patients the PET scans proved to be false negative. These false negative cases included a patient with a pulmonary fibrous tumour (the patient underwent a curative pneumonectomy) and a patient with mantle cell lymphoma (diagnosed 1 year after the PET scan). In one patient the indeterminate solitary pulmonary nodule proved to be true positive at surgery but PET was found to have missed micrometastatic involvement of mediastinal lymph nodes.

Of the patients referred for mediastinal staging, 24 had a positive PET scan result of which 22 were proved to be true positive as shown by pathology in 16 patients and by follow up in six patients; one was found to be false positive (as shown by pathology) and one patient was lost to follow up. Eleven patients had negative scan results which were found to be true negative in 10 patients (as shown by pathology in six patients and by follow up in four: mean time from PET to last chest radiograph or CT scan was 15 months, range 13–17). In one patient the PET scan was found to be false negative (as shown by pathology). In one patient the scan trajectory did not include the mediastinum due to claustrophobia.

Of the patients referred because of distant staging issues, 10 were found to be true positive (as shown by pathology in six patients, follow up in two, and radiology in two). Six patients proved to have a true negative PET scan as shown by follow up in five patients (mean time of follow up 6 months, range 6–6). In one patient the PET result proved to be false negative (bone metastases).
Management changes

In 162 of the 164 cases studied explicit provisional therapeutic plans had been stated before PET. In 103 patients this involved surgery. After PET, surgery was the treatment most commonly abandoned (table 2). PET contributed to a decision to forego surgical treatment in 36 patients (35%; 95% CI 26 to 45) in whom it had been provisionally planned. Of the patients in whom surgery was not the proposed treatment before PET (n=59), seven patients subsequently underwent surgery. In these patients the intended treatment had been observation in four patients, chemotherapy in two patients, and radiotherapy in one. There was a significant change in terms of the “impact” of treatment for the patient, mainly towards a less aggressive approach (surgery→chemotherapy/radiotherapy→observation; p=0.0001). The impact of PET on treatment was divided into major or minor changes as outlined previously. PET was responsible for changes in choice of treatment that were major in 55 patients (66%; 95% CI 55 to 76) and minor in 28 patients (34%; 95% CI 24 to 45).

Post hoc evaluation of treatment choice

The impact of PET on treatment choice was analysed for each scan indication (table 3). According to the attending physician, PET was the most important factor leading to a beneficial change of treatment (TC=5) in 45 of 159 patients (28%; 95% CI 21 to 35) and contributed to such change (TC=4) in 34 (21%; 95% CI 15 to 28).

Some studies have recently addressed the clinical impact of PET. The methodologies and patient spectra were variable, but the reported management changes (65–70%) are uniformly higher than those observed as a by-product in accuracy studies (10–59%). This underlines the fact that management change is multifactorial and does not merely depend on changes of treatment choice for the three different indications were found (p=0.65). Treatment choice ratings after PET scanning indicating malignancy when the suspected lesion was indeed found to be malignant were not different from scans indicating a benign lesion found to be benign (p=0.27). Like diagnostic understanding, the treatment choice ratings were significantly higher for later patients than for early patients (p=0.037).

DISCUSSION

A new test that appears to be more accurate than the standard ones will generate a clinical demand, even if its effect on clinical outcome measures is still unclear. With scarce technology like PET, overconsumption may result precluding general accessibility. Evidence-based guidelines for routine use are therefore needed so that the available scanning capacity can be adjusted to the expected demand. However, guidelines aim at the average patient and may not be applicable in specific situations. In this prospective multicentre before-after study the reported clinical impact of FDG PET as an “add on” technology to solve diagnostic problems in patients with suspected NSCLC was considerable. Clinical compliance with the PET results was high, and PET was reported to have led to beneficial management changes (TC ≥4) in 50% of the patients in the three clinical situations investigated. In addition, a positive influence on diagnostic understanding (DU ≥4) by PET was observed in 84% of the patients. Put in a more conservative way, PET proved to be the key diagnostic tool in one of every four patients referred for PET (DU/TC=5).

Interestingly, we observed an increasing appreciation of PET over time. Even though other explanations may also be valid, individual consultation and feedback as done in our setting is known to improve patient referral patterns. Interpretation of the classification of “important contribution” to treatment choice by PET (TC=4) is not straightforward. It is recognised that, in most clinical situations, decisions are made on the basis of a number of factors. Patient management depends on the preoperative assessment of the probability of disease, which is a joint function of multiple diagnostic indicators such as signs, symptoms and test results together with the effectiveness of the invasive procedures that follow them. This complicates the assessment of the contribution of a single test to a change in patient management. Even though the phrasing of the “contributive” ratings (DU/TC=4) may benefit from accentuation, such positive perceptions may always contain a spectrum of clinical relevance which is difficult to translate into outcome measures. The assessment of the true value of “contributive” rather than directly decisive PET findings (TC=4 v TC=5) is therefore best done in a randomised study design.

Some studies have recently addressed the clinical impact of PET. The methodologies and patient spectra were variable, but the reported management changes (65–70%) are uniformly higher than those observed as a by-product in accuracy studies (10–59%). This underlines the fact that management change is multifactorial and does not merely depend on a single test (such as PET). Alternatively, “clinical value” studies may have overestimated the true clinical contribution of

### Table 2 Treatment changes after PET (TC=4/5, n=78)

<table>
<thead>
<tr>
<th>Treatment change</th>
<th>No. of patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surgery to</td>
<td></td>
</tr>
<tr>
<td>Radiotherapy</td>
<td>6</td>
</tr>
<tr>
<td>Chemotherapy</td>
<td>11</td>
</tr>
<tr>
<td>Observation</td>
<td>18</td>
</tr>
<tr>
<td>Radiotherapy to</td>
<td></td>
</tr>
<tr>
<td>Surgery</td>
<td>1</td>
</tr>
<tr>
<td>Chemotherapy</td>
<td>2</td>
</tr>
<tr>
<td>Observation</td>
<td>3</td>
</tr>
<tr>
<td>Chemotherapy to</td>
<td></td>
</tr>
<tr>
<td>Surgery</td>
<td>2</td>
</tr>
<tr>
<td>Radiotherapy</td>
<td>0</td>
</tr>
<tr>
<td>Observation</td>
<td>2</td>
</tr>
<tr>
<td>Observation to</td>
<td></td>
</tr>
<tr>
<td>Surgery</td>
<td>3</td>
</tr>
<tr>
<td>Radiotherapy</td>
<td>4</td>
</tr>
<tr>
<td>Chemotherapy</td>
<td>1</td>
</tr>
<tr>
<td>Minor changes within:</td>
<td></td>
</tr>
<tr>
<td>Surgery</td>
<td>14</td>
</tr>
<tr>
<td>Radiotherapy</td>
<td>9</td>
</tr>
<tr>
<td>Chemotherapy</td>
<td>2</td>
</tr>
</tbody>
</table>

### Table 3 Impact of PET on patient management and its clinical assessment (treatment choice [TC] ratings as defined in box 1)

<table>
<thead>
<tr>
<th></th>
<th>TC=1</th>
<th>TC=2</th>
<th>TC=3</th>
<th>TC=4</th>
<th>TC=5</th>
<th>Missing</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Radiological abnormality</td>
<td>1</td>
<td>16</td>
<td>42</td>
<td>21</td>
<td>30</td>
<td>2</td>
<td>112</td>
</tr>
<tr>
<td>Mediastinal staging</td>
<td>3</td>
<td>11</td>
<td>10</td>
<td>10</td>
<td>2</td>
<td>36</td>
<td>36</td>
</tr>
<tr>
<td>Distant staging</td>
<td></td>
<td>3</td>
<td>3</td>
<td>5</td>
<td>1</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>Overall</td>
<td>1</td>
<td>22</td>
<td>57</td>
<td>34</td>
<td>45</td>
<td>5</td>
<td>164</td>
</tr>
</tbody>
</table>
PET. Firstly, the clinical impact of a new technology depends on the quality of the previous clinical work-up; poorly performed conventional staging before PET scanning would overestimate its actual value. We therefore made an effort to restrict PET referrals to cases in whom conventional investigations had been performed and had failed. As we have shown, this was the case in most of the patients. Further, a retrospective analysis of the pre-PET work-up showed adherence to internationally accepted guidelines in the majority of patients. Secondly, whether a specific test contributed significantly is a matter of judgement, and thus subject to disagreement, error and imprecise measurement.2 This was, indeed, the case in our study; inconsistencies were identified in 18% of the questionnaire responses. To strengthen the evidence of before-after studies, independent reviewing of the data by experts has been suggested. This has been shown to reduce the presumed benefit of a new technology as assessed by this type of study design.20 However, such findings may also reflect the heterogeneity of daily clinical practice in which patients are actually diagnosed and treated. Thirdly, unconscious bias of the referring clinicians in favour of the new technology may have affected the results. We cannot rule out the possibility that this has occurred, but the opposite may also be true. Even though the sample was not randomly chosen, we found no such effect in the medical records of the cases in which a prolonged follow up was needed and the data were derived from a broad spectrum of hospitals.

The questionnaires used do confirm a distinction between the clinical impact of a test on diagnostic understanding, patient management, and (retrospective) clinical assessment of the appropriateness of these changes. The data clearly show that the perceived benefit of PET scanning consists of altered patient management but, to an even greater extent, of increased diagnostic understanding or confidence in cases where patient management was not altered. In their present form the questionnaires do not allow estimation of the amount of clinical uncertainty. In our opinion, studies such as this may serve to estimate the relative merits of PET for different indications within a specific clinical context. If PET fails to show clinical impact, the presumed indication for PET may be removed from the list, whereas promising results warrant further investigation. Our data do not represent consecutive patients presenting with a similar clinical problem and, as such, our results cannot be extrapolated to imply the routine use of PET in all patients with suspected NSCLC. Estimation of the cost-benefit of such an application requires a direct comparison between patients subjected to PET and conventional work-up. Such a study is currently ongoing in the Netherlands.

In summary, controlled implementation of PET as a “last resort” diagnostic modality improved patient management in at least 25% of clinically problematic cases with suspected NSCLC. The combination of preliminary guidelines, intensive feedback, and prospective monitoring may promote the effective use of scarce technology.

ACKNOWLEDGEMENT
The authors thank A Kalwij and C Karga (secretaries, Clinical PET Center) for collecting all the questionnaires.

Authors’ affiliations
G J Herder, E F Comans, O S Hoekstra, G J Teule, U Joshi, Department of Nuclear Medicine, VU University Medical Center, Amsterdam, The Netherlands
G J Herder, P E Postmus, E F Smit, Department of Pulmonary Medicine, VU University Medical Center, Amsterdam
H van Tinteren, Comprehensive Cancer Center, Amsterdam
O S Hoekstra, Department of Clinical Epidemiology and Biostatistics, VU University Medical Center, Amsterdam

REFERENCES
4 Adams A, Flynn K. Positron emission tomography: descriptive analysis of experience with PET in VA. Technology Assessment Program 10, 1A5.4, Boston, USA, 1998.
Prospective use of serial questionnaires to evaluate the therapeutic efficacy of $^{18}$F-fluorodeoxyglucose (FDG) positron emission tomography (PET) in suspected lung cancer

G J Herder, H van Tinteren, E F Comans, O S Hoekstra, G J Teule, P E Postmus, U Joshi and E F Smit

Thorax 2003 58: 47-51
doi: 10.1136/thorax.58.1.47

Updated information and services can be found at:
http://thorax.bmj.com/content/58/1/47

These include:

References
This article cites 16 articles, 3 of which you can access for free at:
http://thorax.bmj.com/content/58/1/47#BIBL

Email alerting service
Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

Topic Collections
Articles on similar topics can be found in the following collections
- Screening (oncology) (407)
- Radiology (diagnostics) (812)
- Lung cancer (oncology) (670)
- Lung cancer (respiratory medicine) (670)
- Lung neoplasms (608)

Notes

To request permissions go to:
http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to:
http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to:
http://group.bmj.com/subscribe/