Positron emission tomography with 2-[^{18}F]fluoro-2-deoxy-D-glucose: Can it be used to accurately stage the mediastinum in non–small cell lung cancer as an alternative to mediastinoscopy?

Kemp H. Kernstine, MD, PhD

See related article on page 1900.

Last year national health care costs rose 10%. Health care providers and third-party payers are under increasing pressure to reduce costs yet maintain or even improve quality. So can adding an additional test, such as positron emission tomography (PET), to an already expensive evaluation, such as lung cancer assessment, reduce cost and improve quality? It could, if it reduces unnecessary tests and surgery and accurately directs treatment.

Simply, PET is expensive. The high quality, 3-dimensional reconstruction capability, in-computer software systems cost $1.5 to $2.0 million; computed tomography combined systems PET/CT are more than $2.0 million (Sue Ann Halliday, ImageMed Group, LLC, and Barry Siegel, Washington University, St Louis, Mo, personal communications). To provide appropriate space in most medical centers costs an additional $1.5 to $2.0 million. The radiopharmaceutical contrast is also expensive. For the study of lung cancer, most centers use fluorodeoxyglucose F 18 (FDG). It has a brief shelf life, the half-life being 110 minutes. For in-house production, a cyclotron is used that costs $1.0 to $2.0 million. A practical alternative, purchase of commercially made FDG, is approximately $300 to $450 per dose. Thus the initial cost for a complete in-house system is $4 to $6 million dollars. This total does not include the costs of the technicians, part- or full-time physicists, nurses, and nuclear radiologists appropriately trained to run, operate, and maintain the equipment and to read the resultant images. The per-study cost for an ear lobes to pelvis single-patient study in most medical centers is $2800 to $3500, nearly 7 times the cost of a chest/upper abdominal computed tomogram.

The use of PET and the conclusions we make are dependent on two assumptions. First, staging matters. Identifying the state of biologic progression or stage to determine the prognosis and management is important to success in treatment. Second, ideally selected treatment can improve survival and reduce the hindrance of the disease, ultimately reducing the cost of care. Presumably PET should help to discriminate patients who have limited disease, those most effectively treated with surgery.

Generally and not to be completely inclusive, we have learned a number of things about the use of FDG-PET in the clinical evaluation of lung cancer:

1. It is useful in diagnosis and staging,[1] though it neither diagnoses nor excludes the presence of malignancy.
2. It can discriminate less suitable candidates for surgical resection, estimated to be 20% to 25% of potential surgical patients.[2-4]
3. It may provide helpful prognostic information in additional to the staging information.[5-7]

In medical centers that wish to provide complete and accurate staging for their patients and who do not have an accomplished surgeon who can perform mediastinoscopy or thoracoscopy, is PET an alternative to a thoracic surgical investigation?

When reviewing clinical PET research, we must make sure that we understand the differences among the studies. First, the machines used for positron acquisition

From the Division of Cardiothoracic Surgery, The University of Iowa Hospitals and Clinics, Iowa City, Iowa.

Received for publication April 16, 2003; accepted for publication April 21, 2003.

Address for reprints: Kemp H. Kernstine, MD, PhD, The University of Iowa Hospitals, Division of Cardiothoracic Surgery, 200 Hawkins Dr, 1616B-JCP, Iowa City, IA 52242 (E-mail: Kemp-kernstine@uiowa.edu).

J Thorac Cardiovasc Surg 2003;126:1700-3

0022-5223/2003 $30.00 + 0

doi:10.1016/S0022-5223(03)00815-8

have different capabilities. Older and cheaper versions have lower sensitivities. Although reported as PET studies, some studies have not used PET technology at all, instead using coincidence gamma scanning systems. Second, the machines are not truly objective. Interpretation is variable. With greater training and experience, the PET reader develops greater accuracy. Third, the contrast medium in most cases, FDG, has a short half-life, and thus the timing of injection, absorption, and data acquisition may make a difference in the results. Fourth, patient populations differ. Most studies are performed at tertiary care centers and may not represent a generalizable study population. Not all study patients are known to have lung cancer or suspected lung cancer. Some studies exclude patients for a variety of different medical conditions, such as diabetes, or for other individual characteristics. The prevalence of mediastinal cancer varies greatly, with some as low as 5% to 10%, and because of the low sample size it is difficulty to define mediastinal staging accuracy. Fifth, studies differ in how they determine the presence of malignancy in the primary lesion, the hilar or mediastinal nodes, or in the metastases. Some use cytologic examination or follow-up to determine presence of cancer, an option inferior to histologic biopsy. Sixth, the PET interpreters may or many not be aware of the pathologic data. Knowing whether a mass or a node is positive at the time the PET scan is reviewed may influence the result, making the study results less helpful. Seventh, prospective trials are necessary to decrease selection bias and improve our understanding of the modality in question. Finally, many of the studies differ in how they deal with hyperglycemia. Performing FDG-PET on a hyperglycemic patient may dramatically alter the PET result, making it less likely to identify malignancy. In light of all these differences, we should carefully interpret the PET literature, especially when applying the information to our patients.

When it comes to the mediastinum, most studies that were performed before 2000 had relatively few patients who had mediastinal cancer or had undergone thorough histologic biopsy, yet conclusions were made about the accuracy of the PET.[8] Because of this, the research results and conclusions are likely less accurate, and the small sample size results in an inordinately large confidence interval. Poor patient selection, surgical technique, and different PET scanning technology make the results less generalizable. Most studies had fewer than 100 patients. Now with larger trials, some prospective, we can provide helpful information to our colleagues. The technique used to prove the presence of cancer has been variable: observation, bronchoscopy fine-needle aspiration (FNA), computed tomography–directed FNA, computed tomography–directed core needle biopsy, mediastinoscopy–directed FNA, and histologic biopsy by mediastinoscopy, thoracoscopy, or thoracotomy. The success of each technique varies with the skills and experience of the professional performing the biopsy and the location of the lesion or metastasis or node. Also, the statistical techniques were variable. Some counted each node group separately; thus an individual tumor or patient could be overexpressed in the results, resulting in bias. This is difficult, because PET findings rely upon the degree of radioactivity expressed relative to the surrounding area of interest, and nodes or masses adjacent to a particularly high-intensity structure, such as the heart or liver, or to known tumor may be inappropriately regarded as negative for metastases because of the overshadowing effect of the emission system. We must carefully review the various results obtained to see whether they apply to our patients or research question.

In this issue, Gonzalez-Stawinski and colleagues[9] from Duke University report a 5-year analysis of 202 patients with known or suspected lung cancer who underwent PET scanning before mediastinoscopy, 22% of whom had pathologically demonstrable disease. Their primary objective was to compare results of the Duke PET scanning system with those of histologic samples obtained by mediastinoscopy and thoracotomy. Duke is superbly suited to comment on the use of PET in the evaluation of lung cancer patients. It has a long history of PET research, the latest machinery, and highly qualified PET readers, along with exceptional thoracic surgeons and a large patient population from which to draw study participants. Rather than FNA frequently used in previous studies, only histologic tissue samples from mediastinoscopy were compared with the PET results. The mediastinum was evaluated on the PET scans by visual analysis; no attempt was made at software-corrected objectivity such as by using a standard uptake value. The PET readers were blinded to the histologic results. In addition, and what is particularly unique about this trial, they tested their results beyond the traditional pathologic findings by following up their patients for an additional 1 to 2 years to see whether the patients had mediastinal recurrence or died of their disease. No one was unavailable for follow-up. Also uniquely, they assessed the mediastinum as a whole; individual nodes or stations were not separately assessed. Not attempting to unfairly stretch PET technology, the Duke investigators simply asked the question of whether the mediastinum was positive or negative; either by PET, mediastinoscopy, or long-term follow-up. The particularly helpful part of their final results was their finding of 64.4% sensitivity, 77.1% specificity, 44.6% positive predictive value, and 88.3% negative predictive value of PET in the mediastinum.

Their results are similar to my own group's results at the University of Iowa[10] in a similar patient population, a trial size of 237 and 24% with mediastinal disease. We found sensitivity of 68%, specificity of 82%, positive predictive value of 54%, and negative predictive value of 89%. The

results in the two studies are similar. The Duke investigators have concluded, as we did, that relative to mediastinoscopy PET alone is not sufficient to accurately stage the mediastinum and that mediastinoscopy should continue to be performed in potentially operable cases. Like ourselves, they found that PET improved the accuracy of mediastinoscopy. Unlike the Iowa study, they analyzed their mediastinoscopy results with the thoracotomy nodal assessment and followed up their patients with positive PET results but negative pathologic results. Nine of the 32 patients had recurrence within the 2-year follow-up, but none had recurrence in the mediastinum.

There are some inherent problems with the Duke protocol, as with ours:

1. PET centers and PET machines are likely to have different results from Duke. We should expect that at less experienced centers with less expensive PET scanners the positive predictive value and negative predictive value will be worse. Perhaps improved software, better PET machines, and newer contrast media may significantly improve the ability of PET to stage the mediastinum, but with current technology it does not do the job.
2. Only PET data were used to assess the prediction. No information about the patient, computed tomographic data (such as size, location, border, evidence of invasion), other PET data (such as the standard uptake value of the primary, disease elsewhere in the patient, or the standard uptake value was used to predict disease in the mediastinum), or any other clinical results such as information from other scans or tests were included in the analysis. If this information were added to the PET information, it might enhance the prediction of mediastinal cancer.
3. The Duke team's assessment of the mediastinum as a whole is not sufficient for our staging needs. The issue of whether the contralateral or multinodal station disease is very relevant to prognosis and treatment decisions. To be comparable, PET must be able to analyze mediastinal node stations, not necessarily individual nodes, to stage the mediastinum.
4. An inherent problem of the FDG contrast is that inflamed tissues will absorb it. Institutions with a high prevalence of granulomatous or inflammatory mediastinal disease or cases of obstructive malignant processes may have difficulty identifying mediastinal malignancy with PET.
5. For us to make a global statement about the ability of PET to stage, we are excluding a number of patients from the analysis pool, such as those with tracheostomies or uncontrolled diabetes; PET may be worse or better by including these additional patients.
6. Surgeons use the PET results to perform their mediastinoscopy, and thus PET may fare better than mediastinoscopy alone. In the Duke and Iowa studies, PET was not compared with mediastinoscopy alone but with PET plus mediastinoscopy.
7. Pathologic sampling may not be sufficiently accurate to assess involvement of mediastinal nodes.

In a seemingly similar article, different results were reported by Vesselle and colleagues[11] from the University of Washington in September 2002. The University of Washington group retrospectively evaluated their patients with non–small cell lung cancer (NSCLC) who had undergone PET scanning and mediastinoscopy. Relative to the Duke and Iowa studies, the University of Washington group showed that PET scanning was significantly more accurate. The accuracy was 90.7%, sensitivity was 80.9%, specificity was 96%, positive predictive value was 91.9%, and negative predictive value was 90.1%. The questions asked by this group were not about the accuracy of PET staging of the mediastinum. Their primary objective was to evaluate the contribution of PET to the staging of lung cancer in their patients. As a secondary finding, they identified the accuracy of staging the mediastinum. Their study was different from the Duke and Iowa studies. The University of Washington study selected only patients that had a diagnosis of lung cancer, eliminating those with other diagnoses. They eliminated those with primary lesions that were less than 1 cm and eliminated cases that did not appear to have significant nodal pathology according to computed tomography. By doing so, they selected out a group of patients that might have greater PET accuracy. Like the Duke and Iowa Studies, they too found that the PET was helpful at directing the mediastinoscope for biopsy. However, their subselection does not represent the group of patients seeing a thoracic surgeon for resection. The Duke and Iowa studies represent typical patients who would be seen by a thoracic surgeon.

Two meta-analyses have been performed on the subject of the accuracy of FDG-PET to stage the mediastinum.[12,13] At first glance, one might review the large tables from these reviews and, on the basis of their significant cumulative sample sizes, incorrectly conclude that PET accurately stages the mediastinum. This research technique fails from selection bias of the research trials included in the analysis and by the fact that newer techniques or technologies may not be included at the time of the meta-analyses publication. One of the studies, that of Toloza and colleagues, [13] required the publications to have more than 20 patients, publication in a peer-reviewed journal, cytologic examination as admissible, and the availability of raw data for analyses. From the data, receiver operating characteristic analysis was performed on the studies, but no attempt was made to determine whether the science of the research trials was sound and usable. Dwamena and associates[12] made every

effort to perform a complete worldwide analysis of the literature at the time of their publication in 1999. They included only studies in which third-generation scanners or higher were used, patients fasted for at least 4 hours, the interpretation protocol was described, PET interpreters were blinded, there was histologic sampling only, results were reported in sufficient detail for the creation of contingency tables, and established PET criteria were used for the determination of the presence or absence of malignancy. They did not exclude studies in which the primary intent of the study was something other than determining the accuracy of FDG-PET in the mediastinum, the trial appeared to be retrospective, nodal stations or the entire mediastinum were assessed rather than individual nodes, or there were clear criteria for how certain circumstances were accommodated, such as diabetes and tracheostomy. As a result, the conclusions from these meta-analyses too are suspect, and repeated analyses attempting to objectively assess ability of PET to stage the mediastinum may find relatively few articles that possess the qualities necessary to draw helpful conclusions to determine the role of FDG-PET in the evaluation of the patient with NSCLC. The Duke and Iowa articles possess these criteria and have discovered the same results.

Current FDG-PET technology alone does not appear to be sufficient to warrant reliable treatment changes or the avoidance of mediastinoscopy. It is unlikely that the addition of transbronchial, transtracheal, and endoscopic ultrasonographically guided FNA will sufficiently rule out disease relative to the histologic results achieved from mediastinoscopy. However, PET may improve the accuracy of mediastinoscopy, allowing the mediastinoscopist to direct a more thorough biopsy of areas that appear suspect on PET.

It appears that we have added yet another test, and a very expensive one, to the already expensive evaluation of the patient with lung cancer. From the contributions of the Duke and Iowa articles, we cannot eliminate mediastinoscopy in the staging of NSCLS. Seemingly we are increasing the expense, but FDG-PET has helped to discriminate patients with nonsurgically treatable disease and directing them to nonsurgical intervention earlier in the course of their disease, potentially reducing costs and complications while improving outcomes. As the Duke group states in their discussion, we will need to assess the economic impact of this finding. Perhaps with newer PET technology or possibly an alternative tracer, improved software, improved technique of performing PET, and inclusion of additional clinical information, we may be able to exclude mediastinoscopy altogether, at least in a subset of patients. To date we cannot exclude mediastinoscopy in the evaluation of the patient with NSCLC.

References

1. Lowe VJ, Naunheim KS. Positron emission tomography in lung cancer. *Ann Thorac Surg*. 1998;65:1821-9.
2. Pieterman RM, van Putten JW, Meuzelaar JJ, Mooyaart EL, Vaalburg W, Koeter GH, et al. Preoperative staging of non-small-cell lung cancer with positron-emission tomography. *N Engl J Med*. 2000;343: 254-61.
3. Kalff V, Hicks RJ, MacManus MP, Binns DS, McKenzie AF, Ware RE, et al. Clinical impact of (18)F fluorodeoxyglucose positron emission tomography in patients with non-small-cell lung cancer: a prospective study. *J Clin Oncol*. 2001;19:111-8.
4. van Tinteren H, Hoekstra OS, Smit EF, van den Bergh JH, Schreurs AJ, Stallaert RA, et al. Effectiveness of positron emission tomography in the preoperative assessment of patient with suspected non-small-cell lung cancer: the PLUS multicentre randomized trial. *Lancet*. 2002; 359:1388-93.
5. Ahuja V, Coleman RE, Herndon J, Patz EF Jr. The prognostic significance of fluorodeoxyglucose positron emission tomography imaging for patients with nonsmall cell lung carcinoma. *Cancer*. 1998;83:918-24.
6. Vansteenkiste JF, Stroobants SG, Dupont PJ, De Leyn PR, Verbeken EK, Deneffe GJ, et al. Prognostic importance of the standardized uptake value on (18)F-fluoro-2-deoxy-glucose-positron emission tomography scan in non-small cell lung cancer: an analysis of 125 cases. *J Clin Oncol*. 1999;17:3201-6.
7. Patz EF Jr, Connolly J, Herndon J. Prognostic value of FDG-PET imaging after treatment for non–small cell lung cancer. *AJR Am J Roentgenol*. 2000;174:769-74.
8. Vansteenkiste JF, Stroobants SG, De Leyn PR, Dupont PJ, Verschakelen JA, Nackaerts KL, et al. Mediastinal lymph node staging with FDG-PET scan in patients with potentially operable non-small cell lung cancer: a prospective analysis of 50 cases. Leuven Lung Cancer Group. *Chest*. 1997;112:1480-6.
9. Gonzalez-Stawinski GV, Lemaire A, Merchant F, O'Halloran E, Coleman RE, Harpole DH, et al. A comparative analysis of positron emission tomography and mediastinoscopy in staging non–small cell lung cancer. *J Thorac Cardiovasc Surg*. 2003;126:1900-5.
10. Kernstine KH, McLaughlin KA, Menda Y, Rossi NP, Kahn DJ, Bushnell DL, et al. Can FDG-PET reduce the need for mediastinoscopy in potentially resectable nonsmall cell lung cancer? *Ann Thorac Surg*. 2002;73:394-402.
11. Vesselle H, Pugsley JM, Vallieres E, Wood DE. The impact of fluorodeoxyglucose F 18 positron-emission tomography on the surgical staging of non–small cell lung cancer. *J Thorac Cardiovasc Surg*. 2002;124:511-9.
12. Dwamena BA, Sonnad SS, Angolado JF, Wahl RL. Metastases from non-small cell lung cancer: mediastinal staging in the 1990s—meta-analytic comparison of PET and CT. *Radiology*. 1999;213:530-6.
13. Toloza EM, Harpole L, McCrory DC. Noninvasive staging of non–small cell lung cancer: a review of the current evidence. *Chest*. 2003;123(1 Suppl):37S-146S.